

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 20:47:41 ; Search time 6644.53 Seconds
(without alignments)
11088.547 Million cell updates/sec

Title: US-09-434-837-10_COPY_1_1801

Perfect score: 1801

Sequence: 1 atgcagcttcacactccct.....ctatgagtcgagatccga 1801

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_ov:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_da:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_ph:*

24: em_pat:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rnd:*

36: em_htg_mam:*

37: em_htg_vrc:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1765.2	98.0	24595	1 ATACH5	X00493 Agrobacteri
2	1765.2	98.0	24595	6 BD016312	BD016312 Method of
3	1765.2	98.0	24595	6 E00404	E00404 Ti plasmid
4	1765.2	98.0	24595	6 E00546	E00546 DNA fragmen
5	1765.2	98.0	194140	1 AF242881	AF242881 Agrobacte
6	1760.4	97.7	2591	1 T1PMS2	K0254 Ti plasmid
7	1605.2	89.1	206479	1 AB016260	AB016260 Agrobacte
8	1603.6	89.0	9931	1 AE007927	AE007927 Agrobacte
9	1603.6	89.0	10377	1 AE009419	AE009419 Agrobacte
10	1602	89.0	5968	1 AF126446	AF126446 Agrobacte
11	1602	89.0	29524	1 ATU237588	AJ237588 Agrobacte
12	1597.2	88.7	4660	1 AB025110	AB025110 Agrobacte
13	1467.6	81.5	10200	1 ATTA1AAH	X56185 A. tumefacie
14	1467.6	81.5	15463	1 AV083987	U83987 Agrobacteri
15	1464.4	81.3	14960	1 AV083986	U83986 Agrobacteri
16	814	45.2	9577	1 AF126447	AF126447 Agrobacte
17	762.8	42.4	3385	1 AF061780	AF061780 Agrobacte
18	644	35.8	5995	1 ATTU4UX	M61151 Agrobacteri
19	358.2	19.9	5516	1 IS4AAH	M91609 Plasmid pTi
20	258.4	14.3	3677	1 PSE1AAH	M11035 P. syringae
21	254.8	14.1	1937	1 ERW1AM	L33857 Pantoea ag
22	242.6	13.5	4797	1 PSU04358	U04358 Pseudomonas
23	121.6	6.8	2279	6 AR050109	AR050109 Sequence
24	121.6	6.8	2279	6 I24442	I24442 Sequence 8
25	121.6	6.8	21126	1 RIATL	K0313 Integrated
26	121.6	6.8	21126	6 AR050120	AR050120 Sequence
27	121.6	6.8	21126	6 I04770	I04770 Sequence 94
28	121.6	6.8	21126	6 I24453	I24453 Sequence 19
29	66	3.7	217594	1 AP002086	AP002086 Agrobacte
30	65	3.6	167	1 ATTU6787	U16787 Agrobacteri
31	64.4	3.6	1610	1 ARH271050	AJ271050 Agrobacte
32	52.2	2.9	334108	1 AE016857	AE016857 Pseudomon
33	51.4	2.9	1559	8 AK118627	AK118627 Arabidops
34	48.6	2.7	2000	6 AX655393	AX655393 Sequence
35	46	2.6	2000	6 AX655393	AX655393 Sequence
36	43.6	2.4	125020	9 AF429315	AF429315 Homo sapi
37	43.6	2.4	300511	1 AE016775	AE016775 Pseudomon
38	43	2.4	35417	3 AF067936	AF067936 Caenorhab
39	42.4	2.4	125020	9 AF429315	AF429315 Homo sapi
40	41.4	2.3	12733	1 ATTUORF	M60490 Agrobacteri
41	41.4	2.3	163280	2 AC141497	AC141497 Rattus no
42	41.4	2.3	226328	2 AC098150	AC098150 Rattus no
43	41.4	2.3	236068	2 AC121721	AC121721 Rattus no
44	41.4	2.3	240507	2 AC103202	AC103202 Rattus no
45	41.4	2.3	251457	2 AC131820	AC131820 Rattus no

ALIGNMENTS

RESULT 1

LOCUS ATACH5

DEFINITION Agrobacterium tumefaciens Ti plasmid pTi15955-T-DNA region.

ACCESSION X00493

VERSION X00493.1

KEYWORDS octopine synthetase; plasmid; synthetase; terminal repeat; unidentified reading frame.

SOURCE Agrobacterium tumefaciens (Rhizobium radiobacter)

ORGANISM Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.

REFERENCE 1 (bases 1 to 24595)

AUTHORS Barker R.F., Idler K.B., Thompson D.V. and Kemp J.D.

TITLE Nucleotide sequence of the T-DNA region from the Agrobacterium

Pred. No. is the number of results predicted by chance to have a

```
REFERENCE JOURNAL tumefaciens octopine Ti plasmid PT11595
AUTHORS Plant Mol. Biol. 2, 335-350 (1983)
TITLE 2 (bases 602 to 14237)
Glezen,U., De Beuckeleer,M., Seurinck,J., Deboeck,F., De Greve,H.,
Lemmerz,M., Van Montagu,M. and Schell,J.
The complete nucleotide sequence of the Ti-DNA of the Agrobacterium
tumefaciens plasmid pTiAch5
EMBO J. 3 (4), 835-846 (1984)
JOURNAL MEDLINE 84207942
PUBMED 6327292
REFERENCE 3 (bases 1 to 24595)
AUTHORS Turk,S.C., Nester,E.W. and Hooykaas,P.J.
The virA promoter is a host-range determinant in Agrobacterium
tumefaciens
Mol. Microbiol. 7 (5), 719-724 (1993)
JOURNAL MEDLINE 93225814
PUBMED 8469115
REFERENCE 4 (bases 1 to 24595)
AUTHORS Guevara-Garcia,A., Mosqueda-Cano,G., Arguello-Astorga,G.,
Simpson,J. and Herrera-Estrella,L.
Tissue-specific and wound-inducible pattern of expression of the
manopine synthase promoter is determined by the interaction
between positive and negative cis-regulatory elements
Plant J. 4 (3), 495-505 (1993)
JOURNAL MEDLINE 94035196
PUBMED 8220492
FEATURES
source Location/Qualifiers
1. .24595
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:358"
/plasmid="Ti"
909..23782
/note="T-region"
909..932
/note="direct repeat A"
1337
/note="C is T in [2]"
/citation=[2]
1661..2179
/note="open reading frame 1"
/codon_start=1
/translation="MYDGPINIDSSNLQDRRELKLVLIHTENAYRSSAQRSLIAS
QRMNFIINTDVIDPAKDEYKSCRVACCPPTDIPDILNVSLLVYCSFOEMR
RYAQRREYDGVSGAVISTVPPVASEITQTRMLQKVKVQNTSKETHDLDAVIAL
PNIRFKIQISHI"
complement(1736..2113)
/note="open reading frame 2"
/codon_start=1
/translation="CAA25164.1"
/protein_id="CAA25164.1"
/db_xref="GI:39063"
/db_xref="SPTREMBL:04387"
/translation="MCFPACILPKNFLPQPHSLFCYSFGIGMDGNCDAQNAVIES
LQICPSHFKRAINVQORYIEYIKYICWMAARNFALDHFVFSVGNINISINDEI
YPRALSYERTSLRTAISLCVN"
1948
/note="C is T in [2]"
/citation=[2]
2146
/note="T is CG in [2]"
/citation=[2]
complement(2907..3287)
/note="open reading frame 3"
/codon_start=1
/translation="CAA25165.1"
/protein_id="CAA25165.1"
/db_xref="GI:39065"
/db_xref="SPTREMBL:04386"
/translation="MLFNNIGLQTRMFGELSNSTGVVRIIFILNFSFKHWTRLMIPEF
tumefaciens octopine Ti plasmid PT11595
Plant Mol. Biol. 2, 335-350 (1983)
2 (bases 602 to 14237)
Glezen,U., De Beuckeleer,M., Seurinck,J., Deboeck,F., De Greve,H.,
Lemmerz,M., Van Montagu,M. and Schell,J.
The complete nucleotide sequence of the Ti-DNA of the Agrobacterium
tumefaciens plasmid pTiAch5
EMBO J. 3 (4), 835-846 (1984)
84207942
6327292
3 (bases 1 to 24595)
Turk,S.C., Nester,E.W. and Hooykaas,P.J.
The virA promoter is a host-range determinant in Agrobacterium
tumefaciens
Mol. Microbiol. 7 (5), 719-724 (1993)
93225814
8469115
4 (bases 1 to 24595)
Guevara-Garcia,A., Mosqueda-Cano,G., Arguello-Astorga,G.,
Simpson,J. and Herrera-Estrella,L.
Tissue-specific and wound-inducible pattern of expression of the
manopine synthase promoter is determined by the interaction
between positive and negative cis-regulatory elements
Plant J. 4 (3), 495-505 (1993)
94035196
FEATURES
source Location/Qualifiers
1. .24595
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:358"
/plasmid="Ti"
909..23782
/note="T-region"
909..932
/note="direct repeat A"
1337
/note="C is T in [2]"
/citation=[2]
1661..2179
/note="open reading frame 1"
/codon_start=1
/translation="MYDGPINIDSSNLQDRRELKLVLIHTENAYRSSAQRSLIAS
QRMNFIINTDVIDPAKDEYKSCRVACCPPTDIPDILNVSLLVYCSFOEMR
RYAQRREYDGVSGAVISTVPPVASEITQTRMLQKVKVQNTSKETHDLDAVIAL
PNIRFKIQISHI"
complement(1736..2113)
/note="open reading frame 2"
/codon_start=1
/translation="CAA25164.1"
/protein_id="CAA25164.1"
/db_xref="GI:39063"
/db_xref="SPTREMBL:04387"
/translation="MCFPACILPKNFLPQPHSLFCYSFGIGMDGNCDAQNAVIES
LQICPSHFKRAINVQORYIEYIKYICWMAARNFALDHFVFSVGNINISINDEI
YPRALSYERTSLRTAISLCVN"
1948
/note="C is T in [2]"
/citation=[2]
2146
/note="T is CG in [2]"
/citation=[2]
complement(2907..3287)
/note="open reading frame 3"
/codon_start=1
/translation="CAA25165.1"
/protein_id="CAA25165.1"
/db_xref="GI:39065"
/db_xref="SPTREMBL:04386"
/translation="MLFNNIGLQTRMFGELSNSTGVVRIIFILNFSFKHWTRLMIPEF
/translation="MNFADTPLASLDLDWACEEFIKTYGASPOLTEGVEVLOTNNGLLY
LVKGSLSORIHDTLKFKEKEELSTTIKPKAMKMQQSLTFYVAIPQSNVYLCVSN
PEKGFRCRNRPPIYIVAHGMS"
3369..3370
/note="TC is TCG in [2]"
/citation=[2]
3810..3811
/note="AA is A in [2]"
/citation=[2]
complement(4060..5463)
/note="open reading frame 4 gene tms"
/codon_start=1
/translation="CAA25166.1"
/protein_id="CAA25166.1"
/db_xref="GI:39066"
/db_xref="SWISS-PROT:P03868"
/translation="MVAITSLAQSLERLKRKYSCLELVETLIARCEAKSNALLAT
DMDLRSAAKIDRHGNAGVLCGIPCEKANIATVPPTSAATPALINHLPIPSRV
AERLPSGALPGASGMHELSFGITSNAYATGVRPMPNDLIPGSSGVAALASR
LMGGICTDPTGASVRLPALCGVYGRPTLGRYRGRIIPVSPTRTPGIIACQADV
VILDRISGTPERIPPVPLKGRIGLPTTYFDLADVALAETIRILLANGTPEV
EANIPLDELINKASFVVALYEPHALKQYLDVFVTSFSDVIKQIRSDVANINA
OIDGQISAEVELARHSRPRLOATYRVFKLNRLDALFPAIPLVARIPODSSVI
HNGTMIDTFKIVRVNDPSSNAGLPGLSIPVCLTPDLPLVGMEDIDGLADSDRLAIG
GALEBAIGFRYFAGLPN"
5561
/note="T is C in [2]"
/citation=[2]
5809..8076
/note="open reading frame 5 gene tms"
/codon_start=1
/translation="CAA25167.1"
/protein_id="CAA25167.1"
/db_xref="GI:39067"
/db_xref="SPTREMBL:044388"
/translation="MSASPLDNOCDHLPTKAVDLTMVDAKDELDRVSDAFLEBEAS
RGRRIQISTECSAGLACKRLADGRPEISAGKVAVALYIYIKELIGRILESKPW
ARATVGLVAIDAPRCMDPSENOQLQALFLSGKRCAPIDLSHPAIVASISKAGPT
LMPHYENYMKCVTQPTTTLBGAVPFDVPAVGRNMLKSGASGSPFDILVYRPF
DQSDSGRIGFPEDVFKRVAVIGAGISGLVYANLHLAGVDVTTIYESDVGKLT
WSHAFDAPSVVAEMGAMRPPPAFCFLFFLEHYGLSMKPPFNPTGVDYLVYQGV
WMAKQQLPPKLFHRYVNGMRAFLKDGFFHERDVLASPVAILQALSGDIRMAHDSQ
IMLNRRGESFSGSIRIFLGTHPPGSETMSPFDMDLFLKMGIGSGGFGPVEGFTI
ELIRLVINGYENQRCPEGISELPRISAEVYGVVSORIGHOVRAIIOKEXTIK
IRKSGISLBYDVVYVTSGLANQLBHCITCPQNIQOAPVNOAVDNSHTGSKPLM
TERKFWLHLIPLSCVLMDSIACAVALYCLDYESQDPNKGRLVLTSTEDSHKLLAIPD
KERRCLLDLDAISRSFPAPAQHLFPAQADYDQVNIQHDWLTENAGAFKLNRRGDF
VSEELFFQALDPTANDGVYIAGSCSFTGWEAGIOTACNAVCAIINCGGILLKGN
PLEHSWKRNVYRTRN"
complement(6259..6579)
/note="open reading frame 6"
/codon_start=1
/translation="CAA25168.1"
/protein_id="CAA25168.1"
/db_xref="GI:39068"
/db_xref="SPTREMBL:044389"
/translation="MKQPVCHHSSGNASANDRFRRLIRLKRADPSTIGTVKRSV
VVEQVCMGTCRTLQHOVSTISYHKMGHPFGYKPGNAFRAILVQRHMGCSLQAG
SLRD"
6894..6895
/note="TC is CT in [2]"
/citation=[2]
7597
/note="T is C in [2]"
/citation=[2]
7778..8113
/note="open reading frame 7"
/codon_start=1
/translation="CAA25169.1"
/protein_id="CAA25169.1"
/db_xref="GI:39069"
/db_xref="SPTREMBL:044390"
/translation="MLFNNIGLQTRMFGELSNSTGVVRIIFILNFSFKHWTRLMIPEF
```

```

TRHAAVPSQVDMRVLFBRPVTSPVQSLTIVEAFWQRAILSNLTGRLTALISLM
ILQVQL"
conflict
7961
/Note="T is A in [2]"
/citation=(2)
7964
/Note="A is G in [2]"
/citation=(2)
7968. 7969
/Note="CG is GC in [2]"
/citation=(2)
8387
/Note="A is G in [2]"
/citation=(2)
8771. 9493
/Note="open reading frame 8 gene tmr"
CDS
/codon_start=1
/translation="MDHLIFPTCTGKTTTALAQGTGLPVLSDRVQCCQLSTG
SGRPVIELKGTIRLYLDRPLVEGIIAQTAHRLIEVYNHANGLLLEGSTSL
LNCMAANSYSADFRWHIIRKLPDQETPKAKARVKQMLHPAGHSIIQELVYDMN
EPLRLIKEIDGVRAMLPASONQITADWLQLDANMEGKLINGIAOEYFIHARQOE
ORFQVNAAFDGEFEGHFGMY"
conflict
9564
/Note="T is C in [2]"
/citation=(2)
9926. .9927
/Note="TT is T in [2]"
/citation=(2)
9996. .10571
/Note="open reading frame 9"
/codon_start=1
/translation="MDRSMARQGWLPCLSHGKDDQLQGLSELSKYREKQTDILH
TSGSDIINPGSEFLYILDKENYRLCRQWVLVSNASDGLATLLEPSDGYTRQVR
Query Match 98.0%; Score 1765.2; DB 1; Length 24595;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1792; Conservative 0; Mismatches 8; Indels 2; Gaps 2;
QY 1 ATGTCAAGCTTCACCTCTCTTGATTAACAGATGCGATCATCTCCCAACCAAAATGATGAT 60
Db 5809 ATGTCAAGCTTCACCTCTCTTGATTAACAGATGCGATCATCTCCCAACCAAAATGATGAT 5868
QY 61 CTGACAAATGTCGATTAAGCGGATGAATTGACCCGAGGGTTTCCGATGCTTCTTAGAA 120
Db 5869 CTGACAAATGTCGATTAAGCGGATGAATTGACCCGAGGGTTTCCGATGCTTCTTAGAA 5928
QY 121 CGAAGAGCTTCTAGGGGAGAGGATTAACCAAGCTCCACCGAGTGCAGCGCTGGTTA 180
Db 5929 CGAAGAGCTTCTAGGGGAGAGGATTAACCAAGCTCCACCGAGTGCAGCGCTGGTTA 5988
QY 181 GCTTGCAGAAAGGCTGCGATGTCGCTCCCGAGATCTCAGCTGTGGAAGATGACGA 240
Db 5989 GCTTGCAGAAAGGCTGCGATGTCGCTCCCGAGATCTCAGCTGTGGAAGATGACGA 6048
QY 241 GTTCTCTCGCTTATCTATATTGCGAAGAAATTTGCGGCGGATTAATTGAATCGAAA 300
Db 6049 GTTCTCTCGCTTATCTATATTGCGAAGAAATTTGCGGCGGATTAATTGAATCGAAA 6108
QY 301 CTTTGGCGGCGGCAAGTGAAGTGTCTGTTGCCATTCGACTTGGACACCATTTGGATG 360
Db 6109 CTTTGGCGGCGGCAAGTGAAGTGTCTGTTGCCATTCGACTTGGACACCATTTGGATG 6168
QY 361 GATTTCCTCGAAGCAGCACTAATCCAGACCTGTTTTTGTGCTAGCGGTAAAGATGTCGA 420
Db 6169 GATTTCCTCGAAGCAGCACTAATCCAGACCTGTTTTTGTGCTAGCGGTAAAGATGTCGA 6228
```

```

QY 421 CGATTGATCTTAGTATTTGCGGCAATTCATCTTAAGACTGCGCGCTTTCGACC 480
Db 6229 CGATTGATCTTAGTATTTGCGGCAATTCATCTTAAGACTGCGCGCTTTCGACC 6288
QY 481 CTGCAATGCGCGCTGTAAGAGATGCGACAGTAAATGCGTTACCGGGTTTACATACC 540
Db 6289 CTGCAATGCGCGCTGTAAGAGATGCGACAGTAAATGCGTTACCGGGTTTACATACC 6348
QY 541 CTTGAAGGGGCGGTGCTGCTTATGATGCTTATGCTGAAACCTGATGCTGAAGGGT 600
Db 6349 CTTGAAGGGGCGGTGCTGCTTATGATGCTTATGCTGAAACCTGATGCTGAAGGGT 6408
QY 601 TCGGCAAGTTCTCTTCCAAACAATCGACTGCTCTAGAATGCGACGCGTTTTTGAACAA 660
Db 6409 TCGGCAAGTTCTCTTCCAAACAATCGACTGCTCTAGAATGCGACGCGTTTTTGAACAA 6468
QY 661 TGTTCGATAGTGAACGAGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 720
Db 6469 TGTTCGATAGTGAACGAGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 6528
QY 721 GTCATTTGGCGCTGGCATTTTCCGACTGCTGTGTGCAACGAACTGCTTATGCTGGGCTA 780
Db 6529 GTCATTTGGCGCTGGCATTTTCCGACTGCTGTGTGCAACGAACTGCTTATGCTGGGCTA 6588
QY 781 GACGATGTTACAAATATATGAGCAAGTGAATCGTGTGGAAGGCAAGCTTGTGCATAGCT 840
Db 6589 GACGATGTTACAAATATATGAGCAAGTGAATCGTGTGGAAGGCAAGCTTGTGCATAGCT 6648
QY 841 TTCAGGAGCGCTCTAGTGTGCGCGAAATGCGGCGATGCGATTTCTCTGCTGCA 900
Db 6649 TTCAGGAGCGCTCTAGTGTGCGCGAAATGCGGCGATGCGATTTCTCTGCTGCA 6708
QY 901 TTCTGCTTGTTTTCTTCTCTCGAGCGTTACGCGCTGTCTTTCATGAGCGCTTCCCAAT 960
Db 6709 TTCTGCTTGTTTTCTTCTCTCGAGCGTTACGCGCTGTCTTTCATGAGCGCTTCCCAAT 6768
QY 961 CCGCGACAGTGCACACTTACTGTCCTACCAAGGCTCCAAATACATGTTGGAAGCGGG 1020
Db 6769 CCGCGACAGTGCACACTTACTGTCCTACCAAGGCTCCAAATACATGTTGGAAGCGGG 6828
QY 1021 CAGCTCCCAACCGAGCTGTTCATCGGCTTTCACAGCTTGGCGCTGTTTGAAGAC 1080
Db 6829 CAGCTCCCAACCGAGCTGTTCATCGGCTTTCACAGCTTGGCGCTGTTTGAAGAC 6888
QY 1081 GGTTTTCATGAGGAGATATGTGTGCTGCTGCTGTGCTTATTAAGGCTTTGA 1140
Db 6889 GGTTTTCATGAGGAGATATGTGTGCTGCTGCTGTGCTTATTAAGGCTTTGA 6948
QY 1141 TCGAGCACTTAAGTGGGCTCATGACTCTGGCAAAATTTGGTGAACCGTTTGGGAGG 1200
Db 6949 TCGAGCACTTAAGTGGGCTCATGACTCTGGCAAAATTTGGTGAACCGTTTGGGAGG 7008
QY 1201 GAGTCTTCTCTTACAGGATAGAGAGATCTTCTGGGACACATCTCTGCTGCTGAT-AA 1259
Db 7009 GAGTCTTCTCTTACAGGATAGAGAGATCTTCTGGGACACATCTCTGCTGCTGAT-AA 7068
QY 1260 ACAATGAGTTTCTCTATGATTTGGAACCTATTAACCTTAATGGAATAGGATGTCGGG 1319
Db 7069 ACAATGAGTTTCTCTATGATTTGGAACCTATTAACCTTAATGGAATAGGATGTCGGG 7127
QY 1320 GTTGTGCTCAAGTTTGAAGCGGGGTTTATGAGATCTCCGTTGCTGATCAACGAGTA 1379
Db 7128 GTTGTGCTCAAGTTTGAAGCGGGGTTTATGAGATCTCCGTTGCTGATCAACGAGTA 7187
QY 1380 TGAAGAAATCAGCGGATGTCCTTGAAGAAATCTCAGAACTTCCAGTCCGATCCCATC 1439
Db 7188 TGAAGAAATCAGCGGATGTCCTTGAAGAAATCTCAGAACTTCCAGTCCGATCCCATC 7247
QY 1440 TGAAGTGTAAACGCTGTCTGTGAGCAGCGCATATGCCATGTTCAAGTCAAGGCGAT 1499
Db 7248 TGAAGTGTAAACGCTGTGTGTGAGCAGCGCATATGCCATGTTCAAGTCAAGGCGAT 7307
```


QY 1201 GAGTCCTCTCTTCAGGAGATAGAGAGATCTTTCTGGGCAACATCTCTCTGCTGT- AA 1259
DB 7009 GAGTCCTCTCTCTTCAGGAGATAGAGAGATCTTTCTGGGCAACATCTCTCTGCTGTGAA 7068
QY 1260 ACATGAGTTTTTCTCTCATGATTGGACCTATTCAGACTAATGGAAATAGATCTGGGCGG 1319
DB 7069 ACATGAGTTTTTCTCTCATGATTGGACCTATTCAGACTAATGGAAATAGATCTGGGCGG 7127
QY 1320 GTTTGGTCAGTTTTTTTGAAGCGGGTTTATTGAGATCCTCCGCTTGGTCAATCAAGGATA 1379
DB 7128 GTTTGGTCAGTTTTTTTGAAGCGGGTTTATTGAGATCCTCCGCTTGGTCAATCAAGGATA 7187
QY 1380 TGAAGAAATACAGCGAGTGTGCCCTGAAGAAATCTCAGAACTTCCAGCTGCGATC 1439
DB 7188 TGAAGAAATACAGCGAGTGTGCCCTGAAGAAATCTCAGAACTTCCAGCTGCGATC 7247
QY 1440 TGAAGTGTGTTAAGGTGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGCGAT 1499
DB 7248 TGAAGTGTGTTAAGGTGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGCGAT 7307
QY 1500 TCAGAGGAAAGACAAAATAAAGATTAAGGCTTAAGAGGGGATATCTGAATCTTATGA 1559
DB 7308 TCAGAGGAAAGACAAAATAAAGATTAAGGCTTAAGAGGGGATATCTGAATCTTATGA 7367
QY 1560 TAAAGTGTGTGTCATCATCTGACTCGCAATATCCAACTGAGGATTCCTGATGATGCGA 1619
DB 7368 TAAAGTGTGTGTCATCATCTGACTCGCAATATCCAACTGAGGATTCCTGATGATGCGA 7427
QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACCAATATGACAGATC 1679
DB 7428 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACCAATATGACAGATC 7487
QY 1680 GTCAAACTCTCTCTGATGATGAGCAAGAAATCTGTTAGACATATCTCTCCGCTTG 1739
DB 7488 GTCAAACTCTCTCTGATGATGAGCAAGAAATCTGTTAGACATATCTCTCCGCTTG 7547
QY 1740 TGTCTCATGAGAGGATTCGCAAAAGCATGTATTGCTGACTATGATGCGAGATCC 1799
DB 7548 TGTCTCATGAGAGGATTCGCAAAAGCATGTATTGCTGACTATGATGCGAGATCC 7607
QY 1800 GA 1801
DB 7608 GA 7609

RESULT 3
E00404
LOCUS E00404 24595 bp DNA linear PAT 29-SEP-1997
DEFINITION T1 plasmid T-DNA region.
ACCESSION E00404.1 GI:2168687
VERSION E00404.1
KEYWORDS JP 1985156333-A/1.
SOURCE Agrobacterium tumefaciens (Rhizobium radiobacter)
ORGANISM Agrobacterium tumefaciens
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
REFERENCE 1 (bases 1 to 24595)
AUTHORS Garti,E.D., Denitsu,D.S. and Richiyaado,E.B.
TITLE SELECTION USING OPINE SYNTHASE GENE
JOURNAL Patent: JP 1985156333-A 1 16-AUG-1985;
AGURJIENTERIKUSU RES ASSOC LTD
COMMENT OS Agrobacterium tumefaciens
PN JP 1985156333-A/1
PD 16-AUG-1985
PE 14-SEP-1984 JP 1984193841
PR 14-SEP-1983 US 83 532280
PI GARTI EE DAARU, DENISU DABURIYU SATSUTON,
RI RICHIVADO EFU BEIKAA
PC A01H1/00,C12N5/00,C12N15/00,(C12N5/00,C12R1:91); CC
strandedness: Double;
CC topology: Linear;
CC *source: strain=15955;
FH Key Location/Qualifiers

FT repeat_unit 909..1032
FT /note='ROTL (A)'
FT 14060..14083
FT repeat_unit
FT /note='ROTL (B)'
FT 15900..15923
FT repeat_unit
FT /note='ROTL (C)'
FT 23759..23783
FT repeat_unit
FT /note='ROTL (D)'
FT complement(12555..13632)
FT /product='octopine synthase'
FT complement(19147..20128)
FT /product='agropine synthase'
FT 20611..21900
FT /product='agropine synthase'
FT complement(22482..23384)
FT /product='mannopine synthase'.
FEATURES
source Location/Qualifiers
1..24595
/organism='Agrobacterium tumefaciens'
/mol_type='genomic DNA'
/db_xref='taxon:358'
BASE COUNT 6534 a 5509 c 5793 g 6759 t
ORIGIN
Query Match 98.0%; Score 1765.2; DB 6; Length 24595;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1792; Conservative 0; Mismatches 8; Indels 2; Gaps 2;
QY 1 ATGTGAGCTTACCTCTCTGATTAACAGTGGATCATCTCCCAACCAAAATGTGGAT 60
DB 5809 ATGTGAGCTTACCTCTCTGATTAACAGTGGATCATCTCCCAACCAAAATGTGGAT 5868
QY 61 CTGACATGTCATTAAGCGGATGATTAAGACCGGATTCGATGCTTCTTGA 120
DB 5869 CTGACATGTCATTAAGCGGATGATTAAGACCGGATTCGATGCTTCTTGA 5928
QY 121 CGAAGAGCTTCTGAGGAGAGAGATTACTCAAAAGCTCCACGAGTGCAGCGTGGATTA 180
DB 5929 CGAAGAGCTTCTGAGGAGAGAGATTACTCAAAAGCTCCACGAGTGCAGCGTGGATTA 5988
QY 181 GCTTGCAAAAGGCTGCGCATGCTCTCCCGAGATCTCAGCTGTGAAAGTAGCA 240
DB 5989 GCTTGCAAAAGGCTGCGCATGCTCTCCCGAGATCTCAGCTGTGAAAGTAGCA 6048
QY 241 GTTCTCTCCGCTTATATCTATATTGCAAAAGAAATCTGGGCGGATCTTGAATGAAA 300
DB 6049 GTTCTCTCCGCTTATATCTATATTGCAAAAGAAATCTGGGCGGATCTTGAATGAAA 6108
QY 301 CCTTGGGCGGCGCAACAGTGTGCTCGTGGCATGAGCTTGAACCATTTGCATG 360
DB 6109 CCTTGGGCGGCGCAACAGTGTGCTCGTGGCATGAGCTTGAACCATTTGCATG 6168
QY 361 GATTTCCTCGAAGACAACTAATCAAGCCCTGTTTTGCTGAGCGGTAAAGATGCA 420
DB 6169 GATTTCCTCGAAGACAACTAATCAAGCCCTGTTTTGCTGAGCGGTAAAGATGCA 6228
QY 421 CCGATTGATCTTAGTCATTTGCTGATTTCAATCTCTAAGCTGCGGCTTTGCAACC 480
DB 6229 CCGATTGATCTTAGTCATTTGCTGATTTCAATCTCTAAGCTGCGGCTTTGCAACC 6288
QY 481 CTGCCAATGCGCGTGAAGAAATGCGAAGTAATGCGTTACCGGGTTTACATTAAC 540
DB 6289 CTGCCAATGCGCGTGAAGAAATGCGAAGTAATGCGTTACCGGGTTTACATTAAC 6348
QY 541 CTGGAAGGCGCGTGCATTTGATGATGCTTATGATGCGAAACCTGATGCGAAGGCT 600
DB 6349 CTGGAAGGCGCGTGCATTTGATGATGCTTATGATGCGAAACCTGATGCGAAGGCT 6408
QY 601 TCGGCAAGTTCTTTTCAACATGACCTGCTTACGATGCGAGCCGTTTTTGAACAA 660
DB 6409 TCGGCAAGTTCTTTTCAACATGACCTGCTTACGATGCGAGCCGTTTTTGAACAA 6468

QY	661	GGTTCCGATATGAGACCGGATCGGGCTCTTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG	720
Db	6469	TGTTCCGATATGAGACCGGATCGGGCTCTTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG	6528
QY	721	GTCAATGGCGCTGGCATTTCGGAATCGTGGTGGCAACGAATGCTTCATGCTGGGGTA	780
Db	6529	GTCAATGGCGCTGGCATTTCGGAATCGTGGTGGCAACGAATGCTTCATGCTGGGGTA	6588
QY	781	GACGATGTTACAAATATATGAAGCAAGTATCTGTGTGGAGCAAGCTTTGGTCAATGCT	840
Db	6589	GACATGTTACAAATATATGAAGCAAGTATGTTGGAGCAAGCTTTGGTCAATGCT	6648
QY	841	TTCAAGGGACGCTCCGATGATGCTGGCCGAATAGGGGGGAGATGGCATTTCCCTCGCTGCA	900
Db	6649	TTCAAGGGACGCTCCGATGATGCTGGCCGAATAGGGGGGAGATGGCATTTCCCTCGCTGCA	6708
QY	901	TTTCGCTGTATTTTCTTCTCTCGAGCGTTACGGCGCTGTCCTTGATGAGGCGCTTCCCAAT	960
Db	6709	TTTCGCTGTATTTTCTTCTCTCGAGCGTTACGGCGCTGTCCTTGATGAGGCGCTTCCCAAT	6768
QY	961	CCCGGCAAGTTCGACATTAATTGGTCTTACCAAGGCGTCCAAATACATGTGAAAGCCGGG	1020
Db	6769	CCCGGCAAGTTCGACATTAATTGGTCTTACCAAGGCGTCCAAATACATGTGAAAGCCGGG	6828
QY	1021	CAGCTGCCACCGAAGCTGTTCCATCGCGTTTAAACAAGTGGCGGTGTTTAAAGAC	1080
Db	6829	CAGCTGCCACCGAAGCTGTTCCATCGCGTTTAAACAAGTGGCGGTGTTTAAAGAC	6888
QY	1081	GCTTTTATGAGACGAGATATTGTGTGGCTTCGCTGCTGCTATATCTCAGAGCCCTGAA	1140
Db	6889	GCTTTTATGAGACGAGATATTGTGTGGCTTCGCTGCTGCTATATCTCAGAGCCCTGAA	6948
QY	1141	TCAGGACACATTAAGTGGGCTCATGATCTCTGGCAAAATTTGGCTGAACCGTTTGGAGG	1200
Db	6949	TCAGGAGACATTAAGTGGGCTCATGATCTCTGGCAAAATTTGGCTGAACCGTTTGGAGG	7008
QY	1201	GAGTCTCTTCTCTTCAAGGATAGAGAGATCTTTCTGGGCAACAATCTCTCGTGTGT-AA	1259
Db	7009	GAGTCTCTTCTCTTCAAGGATAGAGAGATCTTTCTGGGCAACAATCTCTCGTGTGTAA	7068
QY	1260	ACATGGAATTTTCCCATGATTTGGGACCTATTCGAAGCTTAAAGGAATAGATCTGGCGGG	1319
Db	7069	ACATGGAATTTTCCCATGATTTGGGACCTATTCGAAGCTTAAAGGAATAGATCTGGCG-GG	7127
QY	1320	GTTTGGTCAAGTTTGAAGCGGGTTTATTGAATCTCCGCTTGGTGCATCAACGATA	1379
Db	7128	GTTTGGTCAAGTTTGAAGCGGGTTTATTGAATCTCCGCTTGGTGCATCAACGATA	7187
QY	1380	TGAAGAAATCAGCGGATGTGCTCCCTGAAGGAATCTCAGAACTTCCAGTCCGGATTCGCATC	1439
Db	7188	TGAAGAAATCAGCGGATGTGCTCCCTGAAGGAATCTCAGAACTTCCAGTCCGGATTCGCATC	7247
QY	1440	TGAAGTGGTTAAAGGTGTGCTGTGAGCGAGGCAATATGCCATGTTCAAGTCAGGGCGAT	1499
Db	7248	TGAAGTGGTTAAAGGTGTGCTGTGAGCGAGGCAATATGCCATGTTCAAGTCAGGGCGAT	7307
QY	1500	TCAGAGGAAAGACAAATAATGAATTAAGCTTAAGGCTTAAGACCGGATATCTGAATTTATGA	1559
Db	7308	TCAGAGGAAAGACAAATAATGAATTAAGGCTTAAGGCGGGATATCTGAATTTATGA	7367
QY	1560	TAAAGTGTGTGCATCTGGAATCTGCAAAATATCCATCAGGCAATTCCTGACATGCGGA	1619
Db	7368	TAAAGTGTGTGCATCTGGAATCTGCAAAATATCCATCAGGCAATTCCTGACATGCGGA	7427
QY	1620	TACCAATATTTTTCAGGACACAGTGAACCAAGCGGTATTAACAGCCATATGACAGATC	1679
Db	7428	TACCAATATTTTTCAGGACACAGTGAACCAAGCGGTATTAACAGCCATATGACAGATC	7487
QY	1680	GTCAAAATCTTCTCGATGATGAAACGAAATTCGTGTTAGACCAATATCTCCGCTTGG	1739
Db	7488	GTCAAAATCTTCTCGATGATGAAACGAAATTCGTGTTAGACCAATATCTCCGCTTGG	7547
QY	1740	TGTTCTCATGACGGGATCGAAAACAGTGAATTTGCTCGACATATGATGCTGACAGATTC	1799

QY	Db	1800 GA 1801	7548 TGTCCTCATGATGAGCGGATCCCAAAAGCATGTATTGCTTGCACTAGATGCGACGATCC	7607
QY	Db	7608 GA 7609		
RESULT 4				
LOCUS E00546 24595 bp DNA linear PAT 29-SEP-1997				
DEFINITION DNA fragment of T-DNA of Ti-plasmid, pRI15955.				
ACCESSION E00546				
VERSION E00546.1 GI:2168825				
KEYWORDS UP 1985256383-A/1.				
SOURCE Agrobacterium tumefaciens (Rhizobium radiobacter)				
ORGANISM Agrobacterium tumefaciens				
REFERENCE Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;				
JOURNAL Rhizobiaceae; Rhizobium; Agrobacterium group; Agrobacterium.				
1 (bases 1 to 24595)				
Richiyado, E.B. and Jiyon, D.K.				
PROMOTOR OF OCTOPINE T-DNA				
Patent: JP 1985256383-A 1 18-DEC-1985;				
AGURIJENREIKUSU RES ASSOC LTD				
OS Agrobacterium tumefaciens				
PN JP 1985256383-A/1				
PD 18-DEC-1985				
PF 19-NOV-1984 JP 1984244307				
PR 18-NOV-1983 US 83 553786				
PI RICHIVADO EPV BEIKAA, JIYON DEI KENPU				
PC C12N15/00, A01H1/00, C12N1/20, C12N5/00//C07H21/04, (C12N1/20, PC				
C12R1:01)				
PC (C12N5/00, C12R1:91);				
CC strandedness: Single;				
CC topology: Linear;				
CC hypothetical: No;				
CC anti-sense: No;				
FH key Location/Qualifiers				
FH CDS				
FT >1..<24595				
FT /gene='tms'				
FT >1..<24595				
FT /gene='tmr'				
FT >1..<24595				
FT /gene='tml'				
FT >1..<24595				
FT /gene='ocs'				
FT /product='octopine synthetase'.				
FEATURES				
Source Location/Qualifiers				
1..24595				
/organism='Agrobacterium tumefaciens'				
/mol_type='genomic DNA'				
/db_xref='taxon:358'				
BASE COUNT 6534 a 5509 c 5793 g 6759 t				
ORIGIN				
Query Match 98.0%; Score 1765.2; DB 6; Length 24595;				
Best Local Similarity 99.4%; Pred. No. 0;				
Matches 1792; Conservative 0; Mismatches 8; Indels 2; Gaps 2;				
QY	Db	1	ATGTCAGCTTCACCTCTCTCTGTAATACACAGCGCATGATCTCCCAACCAAAATGGTGGAT	60
Db	5809	ATGTCACCTTCACCTCTCTCTGTAATACACAGCGCATATCTCCCAACCAAAATGGTGGAT	5868	
QY	61	CTGACAAATGGCGATAAGCGGATGATTTGGAACCGACGAGGTTTCCGATGCTTCTTAGAA	120	
Db	5869	CTGACAAATGGCGATAAGCGGATGATTTGGAACCGACGAGGTTTCCGATGCTTCTTAGAA	5928	
QY	121	CGAAGACCTTTAGAGGGAAGAGGATTAATCTCAAGCTTCCACCGATGCGACGCTGGGTTA	180	
Db	5929	CGAAGACCTTTAGAGGGAAGAGGATTAATCTCAAGCTTCCACCGATGCGACGCTGGGTTA	5988	
QY	181	GCTTGCAAAAGGCTGGCGCATGTCGCTTCCCGCAGATCTCAGCTGGTGGAAAAGTACGA	240	

Db 5989 GCTTGGCAAAAGGCGCGCATGTCGCTCCCGAGATCTAGCTGTGAAAGGTAGCA 6048
 Qy 241 GTTCTCCCGCTTATATCTAATATGGCAAGAAATTTGGGGCGGATACTTGAATCGAA 300
 Db 6049 GTTCTCCCGCTTATATCTAATATGGCAAGAAATTTGGGGCGGATACTTGAATCGAA 6108
 Qy 301 CTTTGGGCGGGGCAAGTAGTGTCTGTGTCATCTGCACTTGAACACCATTTGCATG 360
 Db 6109 CTTTGGGCGGGGCAAGTAGTGTCTGTGTCATCTGCACTTGAACACCATTTGCATG 6168
 Qy 361 GAATTTCTCGAAGACAACTAATTCAGCCCTGTTTTTGTGAGCGGTAAAGATGCA 420
 Db 6169 GAATTTCTCGAAGACAACTAATTCAGCCCTGTTTTTGTGAGCGGTAAAGATGCA 6228
 Qy 421 CCGATTATCTTAGTCAATTTCTGTGCGCAATTCATCTCTAAGACTGCCGCTTTCGACC 480
 Db 6229 CCGATTATCTTAGTCAATTTCTGTGCGCAATTCATCTCTAAGACTGCCGCTTTCGACC 6288
 Qy 481 CTGCGCAATGCGCGTGTAGCAAGATGAGACGATGAATGCGTTACCGGGTTTACCATACC 540
 Db 6289 CTGCGCAATGCGCGTGTAGCAAGATGAGACGATGAATGCGTTACCGGGTTTACCATACC 6348
 Qy 541 CTGGAAGGGGCGCGCATTTGACATGTAGCTTATGCTGAAACCTGATGCTGAAGGCT 600
 Db 6349 CTGGAAGGGGCGCGCATTTGACATGTAGCTTATGCTGAAACCTGATGCTGAAGGCT 6408
 Qy 601 TCGGCAAGGTTCTTTCCAAATGCACTTGCTCTACAGCTGACAGCCGTTTTTGAACAA 660
 Db 6409 TCGGCAAGGTTCTTTCCAAATGCACTTGCTCTACAGCTGACAGCCGTTTTTGAACAA 6468
 Qy 661 TGTTCGATAGTGAAGGATCGGCTTCTTCCGAGGATGTTCTTAACCGGAAGTGGCG 720
 Db 6469 TGTTCGATAGTGAAGGATCGGCTTCTTCCGAGGATGTTCTTAACCGGAAGTGGCG 6528
 Qy 721 GTGATTGGCGCTGGCAATTTCCGAGCTGTGTGGCAACGAACTGCTTCACTGCTGGGTA 780
 Db 6529 GTGATTGGCGCTGGCAATTTCCGAGCTGTGTGGCAACGAACTGCTTCACTGCTGGGTA 6588
 Qy 781 GACGATGTTCAATATATGAAGCAAGTATGCTGTGGAGCAAGCTTTGCTCATGCT 840
 Db 6589 GACGATGTTCAATATATGAAGCAAGTATGCTGTGGAGCAAGCTTTGCTCATGCT 6648
 Qy 841 TTGAGGAGCGCTCTAGTGTGTGGCGGAATGGGGCGAATGCGATTCTCTCCGCTGCA 900
 Db 6649 TTGAGGAGCGCTCTAGTGTGTGGCGGAATGGGGCGAATGCGATTCTCTCCGCTGCA 6708
 Qy 901 TTTGCTGTTTTTCTTCTCTGAGCGTGTAGCGCTGTCTGATGAGGCGTTCCCAAT 960
 Db 6709 TTTGCTGTTTTTCTTCTCTGAGCGTGTAGCGCTGTCTGATGAGGCGTTCCCAAT 6768
 Qy 961 CCGGCGACAGTCGACACTTACTGTTGCTTACCAAGCGCTTCAATATCATGTGAAAAGCGGG 1020
 Db 6769 CCGGCGACAGTCGACACTTACTGTTGCTTACCAAGCGCTTCAATATCATGTGAAAAGCGGG 6828
 Qy 1021 CACTGCGACCGAAGCTTCTCAATCGGCTTAAACCGTTGGGTGGTCTTGAAGGAC 1080
 Db 6829 CACTGCGACCGAAGCTTCTCAATCGGCTTAAACCGTTGGGTGGTCTTGAAGGAC 6888
 Qy 1081 GGTTCATGAGGAGATATGTGTGGCTTCCGCTGCTATTAATCTCAGGCGCTTGA 1140
 Db 6889 GGTTCATGAGGAGATATGTGTGGCTTCCGCTGCTATTAATCTCAGGCGCTTGA 6948
 Qy 1141 TCAAGACACTATAGTGGGCTCATGACTCTGCGAAATTTGGCTGAACCGTTTGGGAGG 1200
 Db 6949 TCAAGACACTATAGTGGGCTCATGACTCTGCGAAATTTGGCTGAACCGTTTGGGAGG 7008
 Qy 1201 GAGTCCCTCTCTCAGGAGATAGAGAGATCTTTCTGGGCAACATCTCTGCTGTGT -AA 1259
 Db 7009 GAGTCCCTCTCTCAGGAGATAGAGAGATCTTTCTGGGCAACATCTCTGCTGTGTGTGAA 7068
 Qy 1260 ACATGAGTTTCTCTCATGATGAGGACTATTCAGGCTAATGGAATAGGATCTGGCGG 1319

Db 7069 ACATGAGTTTCTCTCATGATTTGGGACCTATTCAGCTAATGGGAATAGGATCTGGC -GG 7127
 Qy 1320 GTTGGTCCAGTTTGTGAAGCGGGTTTATTTGATCTCCCGTGTGATCAACGGAT 1379
 Db 7128 GTTGGTCCAGTTTGTGAAGCGGGTTTATTTGATCTCCCGTGTGATCAACGGAT 7187
 Qy 1380 TGAAGAAAATCAGCGATGTGCCCTGAAGAAATCTCAGAACTTCCAGTTCGATTC 1439
 Db 7188 TGAAGAAAATCAGCGGATGTGCCCTGAAGAAATCTCAGAACTTCCAGTTCGATTC 7247
 Qy 1440 TGAAGTGTAAACGTTGTCTGTGTGAGCCAGCATATGCAATGTTCAAGTCAGGCGAT 1499
 Db 7248 TGAAGTGTAAACGTTGTCTGTGTGAGCCAGCATATGCAATGTTCAAGTCAGGCGAT 7307
 Qy 1500 TCGAAGGAAAAGACAAAATTAAGTAAAGCTTAAGACCGGATATCTGAATTTATGA 1559
 Db 7308 TCGAAGGAAAAGACAAAATTAAGTAAAGCTTAAGACCGGATATCTGAATTTATGA 7367
 Qy 1560 TAAAGTGTGTGATCATCTGCACTCGCAATATCCAACTCAGGCAATTCGCTGATGCGA 1619
 Db 7368 TAAAGTGTGTGATCATCTGCACTCGCAATATCCAACTCAGGCAATTCGCTGATGCGA 7427
 Qy 1620 TACCAATATTTTTCAGGCAACAGTGAACCAAGCGTTGATTAACAGCCATATGACAGATC 1679
 Db 7428 TACCAATATTTTTCAGGCAACAGTGAACCAAGCGTTGATTAACAGCCATATGACAGATC 7487
 Qy 1680 GTCAAAATCTTCTCTGATGATGAAAGAAATTTCTGTTAGACATATCTCCCGTCTTG 1739
 Db 7488 GTCAAAATCTTCTCTGATGATGAAAGAAATTTCTGTTAGACATATCTCCCGTCTTG 7547
 Qy 1740 TGTCTCATGAGACGGGATCGCAAAAGCAGTGTATGCTCGACTATGAGTCCGAGATCC 1799
 Db 7548 TGTCTCATGAGACGGGATCGCAAAAGCAGTGTATGCTCGACTATGAGTCCGAGATCC 7607
 Qy 1800 GA 1801
 Db 7608 GA 7609

RESULT 5
 AF242881
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AF242881 194140 bp DNA circular BCT 26-JUN-2000
 Agrobacterium tumefaciens octopine-type Ti plasmid, complete
 sequence.
 AF242881 AF034854 AF034855 AF035413 AF035773 AF039887 AF039888
 AF117204 AF006976 M1480 M14762 M80605 M80607 U19620 U43674 U43675
 U48718 U60011 X04784
 AF242881.1 GI:8572673

Agrobacterium tumefaciens (Rhizobium radiobacter)
 Agrobacterium tumefaciens
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 1 (bases 193234 to 194140; 1 to 23692)
 Barker R.F., Idler K.B., Thompson D.V. and Kemp J.D.
 Nucleotide sequence of the T-DNA region from the Agrobacterium
 tumefaciens octopine Ti plasmid pTi1595
 Plant Mol. Biol. 2, 335-350 (1983)
 2 (bases 166443 to 168193)
 original sequence presented in GenBank Accession Number X00493
 Yanofsky, M.F. and Nester, E.W.
 Molecular characterization of a host-range-determining locus from
 Agrobacterium tumefaciens
 J. Bacteriol. 168 (1), 244-250 (1986)
 MEDLINE 87008387
 PUBMED 3759904
 JOURNAL 3 (bases 168089 to 173951)
 Yanofsky, M.F., Potter, S.G., Young, C., Albright, L.M., Gordon, M.P.
 and Nester, E.W.
 The vird operon of Agrobacterium tumefaciens encodes a
 site-specific endonuclease
 Cell 47 (3), 471-477 (1986)
 JOURNAL MEDLINE 87028239

- PUBMED 3021341
REFERENCE 4 (bases 175278 to 177719)
AUTHORS Winans,S.C., Allena,P., Stachel,S.E., McBride,K.E. and Neeter,E.W.
TITLE Characterization of the vire operon of the Agrobacterium Ti plasmid pTiA6
- JOURNAL Nucleic Acids Res. 15 (2), 825-837 (1987)
MEDLINE 87146396
PUBMED 3547330
REFERENCE 5 (bases 50125 to 54471)
AUTHORS Habeb,L.F., Wang,L. and Winans,S.C.
TITLE Transcription of the octopine catabolism operon of the Agrobacterium tumor-inducing plasmid pTiA6 is activated by a tyrosine-type regulatory protein
- JOURNAL Mol. Plant Microbe Interact. 4 (4), 379-385 (1991)
MEDLINE 92190631
PUBMED 1799699
REFERENCE 6 (bases 50125 to 54471)
AUTHORS Valdivia,R.H., Wang,L. and Winans,S.C.
TITLE Characterization of a putative periplasmic transport system for octopine accumulation encoded by Agrobacterium tumefaciens Ti plasmid pTiA6
- JOURNAL J. Bacteriol. 173 (20), 6398-6405 (1991)
MEDLINE 92011411
PUBMED 1655707
REFERENCE 7 (bases 68024 to 76460)
AUTHORS Kim,K.S. and Farrand,S.K.
TITLE Ti plasmid-encoded genes responsible for catabolism of the crown gall opine mannopine by Agrobacterium tumefaciens are homologs of the T-region genes responsible for synthesis of this opine by the plant tumor
- JOURNAL J. Bacteriol. 178 (11), 3275-3284 (1996)
MEDLINE 96236046
PUBMED 8655509
REFERENCE 8 (bases 39167 to 45466)
AUTHORS Fuqua,C. and Winans,S.C.
TITLE Localization of OcrR-activated and TrfA-activated promoters that express two ABC-type permeases and the trfA gene of Ti plasmid pTiR10
- JOURNAL Mol. Microbiol. 20 (6), 1199-1210 (1996)
MEDLINE 96405643
PUBMED 8809772
REFERENCE 9 (bases 22855 to 39243; 96353 to 111409)
AUTHORS Alt-Worbe,J., Styrker,J.L., Fuqua,C., Li,P.L., Farrand,S.K. and Winans,S.C.
TITLE The conjugal transfer system of Agrobacterium tumefaciens octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid vir genes
- JOURNAL J. Bacteriol. 178 (14), 4248-4257 (1996)
MEDLINE 96312368
PUBMED 8763954
REFERENCE 10 (bases 54312 to 62806)
AUTHORS Oger,P., Kim,K.S., Sackett,R.L., Piper,K.R. and Farrand,S.K.
TITLE Octopine-type Ti plasmid code for a mannopine-inducible dominant-negative allele of trfA, the quorum-sensing activator that regulates Ti plasmid conjugal transfer
- JOURNAL Mol. Microbiol. 27 (2), 277-288 (1998)
MEDLINE 98143415
PUBMED 9484884
REFERENCE 11 (bases 133963 to 138972)
AUTHORS Kallogeraki,V.S. and Winans,S.C.
TITLE Wound-released chemical signals may elicit multiple responses from an Agrobacterium tumefaciens strain containing an octopine-type Ti plasmid
- JOURNAL J. Bacteriol. 180 (21), 5660-5667 (1998)
MEDLINE 99009000
PUBMED 9791116
REFERENCE 12 (bases 76455 to 98723)
AUTHORS Lyi,S.M., Jafari,S. and Winans,S.C.
TITLE Mannopinic acid and agropinic acid catabolism region of the octopine-type Ti plasmid pTi15955
- JOURNAL Mol. Microbiol. 31 (1), 339-347 (1999)
MEDLINE 99141607
PUBMED 9987134
- REFERENCE 13 (bases 1 to 194140)
AUTHORS Winans,S.C., Zhu,J., Oger,P.M., Schrammeyer,B., Hooykaas,P.J. and Farrand,S.K.
TITLE Octopine-type Ti plasmid sequence
- JOURNAL Unpublished
REFERENCE 14 (bases 68024 to 76460)
AUTHORS Kim,K.-S.
TITLE Direct Submission
- JOURNAL Submitted (10-JAN-1995) Kun-Soo Kim, Microbiology, University of Illinois at Urbana-Champaign, 1201 W. Gregory Dr., Urbana, IL 61801, USA
- REFERENCE 15 (bases 96353 to 111409)
AUTHORS Winans,S.C.
TITLE Direct Submission
- JOURNAL Submitted (20-DEC-1995) Stephen C. Winans, Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE 16 (bases 39167 to 45466)
AUTHORS Fuqua,C.
TITLE Direct Submission
- JOURNAL Submitted (07-FEB-1996) Clay Fuqua, Biology, Trinity University, 715 Stadium Dr., San Antonio, TX 78212, USA
- REFERENCE 17 (bases 54312 to 62806)
AUTHORS Kim,K.-S., Sackett,R.L. and Farrand,S.K.
TITLE Direct Submission
- JOURNAL Submitted (04-JUN-1996) Crop Sciences, University of Illinois at Urbana-Champaign, 240 ERML, 1201 W. Gregory Dr., Urbana, IL 61801, USA
- REFERENCE 18 (bases 186239 to 193239)
AUTHORS Zhu,J. and Winans,S.C.
TITLE Direct Submission
- JOURNAL Submitted (17-NOV-1997) Section of Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE 19 (bases 76455 to 98723)
AUTHORS Winans,S.C. and Lyi,S.
TITLE Direct Submission
- JOURNAL Submitted (20-NOV-1997) Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE 20 (bases 170209 to 175283)
AUTHORS Zhu,J. and Winans,S.C.
TITLE Direct Submission
- JOURNAL Submitted (25-NOV-1997) Section of Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE 21 (bases 133963 to 138972; 176911 to 186244)
AUTHORS Styrker,J.L., Mantlis,N.J., Kallogeraki,V.S. and Winans,S.C.
TITLE Direct Submission
- JOURNAL Submitted (23-DEC-1997) Microbiology, Cornell University, Ithaca, NY 14853, USA
- REFERENCE 22 (bases 22855 to 39243)
AUTHORS Winans,S.C.
TITLE Direct Submission
- JOURNAL Submitted (03-AUG-1998) Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE 23 (bases 62801 to 66985)
AUTHORS Oger,P.M.
TITLE Direct Submission
- JOURNAL Submitted (29-DEC-1998) Crop Sciences, University of Illinois at Urbana-Champaign, 361 ERML, 1201 W. Gregory Dr., Urbana, IL 61801, USA
- REFERENCE 24 (bases 110822 to 133968)
AUTHORS Zhu,J. and Winans,S.C.
TITLE Direct Submission
- JOURNAL Submitted (18-AUG-1999) Section of Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE 25 (bases 1 to 194140)
AUTHORS Zhu,J., Oger,P.M., Schrammeyer,B., Hooykaas,P.J., Farrand,S.K. and Winans,S.C.
TITLE Direct Submission
- JOURNAL Submitted (07-MAR-2000) Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- COMMENT On or before Jun 20 2000 this sequence version replaced
gi:3377769, gi:1125729, gi:1381799, gi:797330, gi:2665704,
gi:2749889, gi:5758260, gi:2773254, gi:2773255, gi:2773256,
gi:154820, gi:154827, gi:2662524, gi:39122, gi:2687604, gi:154767,

gi:154773, gi:6689862.
Location/Qualifiers
source 1..194140
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:358"
/plasmid="T1"
/note="extrachromosomal
octopline-lypse"
13244..13618
source /organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:358"
/insertion_seq="ybd"
13636..14181
source /organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:358"
/insertion_seq="ybe"
/note="similar to 1866"
14183..14666
source /organism="Agrobacterium tumefaciens"

Query Match 98.0%; Score 1765.2; DB 1; Length 194140;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1792; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 1 ATGTGACCTTCACTCTCTTGTATACCAAGTGCATATCTCCACCAAAATGTGGAT 60
DB 4902 ATGTGACCTTCACTCTCTTGTATACCAAGTGCATATCTCCACCAAAATGTGGAT 4961
QY 61 CTGACATAGTGTGATAGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 4962 CTGACATAGTGTGATAGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 5021
QY 121 CGAAGACCTTCTAGGGGAGAGGATTAACAAGCTCAACGAGTGCAGCGCTGAGTTA 180
DB 5022 CGAAGACCTTCTAGGGGAGAGGATTAACAAGCTCAACGAGTGCAGCGCTGAGTTA 5081
QY 181 GCTTGCAAAAGCTGCGCATGCTGCTTCCCGAGATCTGAGTGTGAAAAGTGA 240
DB 5082 GCTTGCAAAAGCTGCGCATGCTGCTTCCCGAGATCTGAGTGTGAAAAGTGA 5141
QY 241 GTTCTCCGCTTATATCTATATTTGGCAAGAAATCTGGGGGGGATTAATGCA 300
DB 5142 GTTCTCCGCTTATATCTATATTTGGCAAGAAATCTGGGGGGGATTAATGCA 5201
QY 301 CCTTGGGCGGGCAACAGTGAAGTGTCTGCTGCCATCGACTTGAACACATTTGCATG 360
DB 5202 CCTTGGGCGGGCAACAGTGAAGTGTCTGCTGCCATCGACTTGAACACATTTGCATG 5261
QY 361 GATTTCCTCGAAGACAACTAATCCAGCCCTGTTTTTCTGAGCGGTAAGATGTGA 420
DB 5262 GATTTCCTCGAAGACAACTAATCCAGCCCTGTTTTTCTGAGCGGTAAGATGTGA 5321
QY 421 CGCATGATCTTAGTCAATTCGTGGCCATTTCAATCTCTAAGAATGCGCGCTTTGCAAC 480
DB 5322 CGCATGATCTTAGTCAATTCGTGGCCATTTCAATCTCTAAGAATGCGCGCTTTGCAAC 5381
QY 481 CTGCCAATGCGCGTGAAGAAATGACAGATGAATGCAATGCGGTATTCATCAACC 540
DB 5382 CTGCCAATGCGCGTGAAGAAATGACAGATGAATGCAATGCGGTATTCATCAACC 5441
QY 541 CTGGAAGGGGCGGTCATTTGACATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 5442 CTGGAAGGGGCGGTCATTTGACATGATGATGATGATGATGATGATGATGATGATGAT 5501
QY 601 TCGGCAAGTCTCTTCAACAATGCAATGCTCTGACAGTGCAGACCGTTTTTGGACCA 660
DB 5502 TCGGCAAGTCTCTTCAACAATGCAATGCTCTGACAGTGCAGACCGTTTTTGGACCA 5561
QY 661 TGTTCCTATGATGACGATCGGCTTCTTCCGAGAGATGCTTAAGCCGAAGTGGCG 720

DB 5562 TGTTCCTATGATGACGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 5621
QY 721 GTCAATGGGCTGGCATTTCCGCACTGATGGGCAAGAACTGCTTCAATGCTGGAGTA 780
DB 5622 GTCAATGGGCTGGCATTTCCGCACTGATGGGCAAGAACTGCTTCAATGCTGGAGTA 5681
QY 781 GAGATGTTACATATATATGAAGCAAGTGAATGATGATGATGATGATGATGATGATGAT 840
DB 5682 GAGATGTTACATATATATGAAGCAAGTGAATGATGATGATGATGATGATGATGATGAT 5741
QY 841 TTCAGGAGCGCTCTAATGCTGTGAGCGCAATGAGGCGCATGATGATGATGATGATGAT 900
DB 5742 TTCAGGAGCGCTCTAATGCTGTGAGCGCAATGAGGCGCATGATGATGATGATGATGAT 5801
QY 901 TTCTGCTGTTTTTCTTCCGAGCGCTTACGCGCTGCTTCTGATGAGCGCTTCCCAAT 960
DB 5802 TTCTGCTGTTTTTCTTCCGAGCGCTTACGCGCTGCTTCTGATGAGCGCTTCCCAAT 5861
QY 961 CCGGCAAGTGCACACTTACTTGTCTACCAAGGCGTCCAAATACATGTTGAAGCGGG 1020
DB 5862 CCGGCAAGTGCACACTTACTTGTCTACCAAGGCGTCCAAATACATGTTGAAGCGGG 5921
QY 1021 CAGCTGCCAAGCGAGCTGTCATGCGGTTTCAACAGGTTGGCGCTTCTTGAAGAC 1080
DB 5922 CAGCTGCCAAGCGAGCTGTCATGCGGTTTCAACAGGTTGGCGCTTCTTGAAGAC 5981
QY 1081 GGTTCCTATGAGGAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 5982 GGTTCCTATGAGGAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 6041
QY 1141 TCAGACACATTAAGGAGGCTCATGACTCTGCAAAATTTGGTGAACGTTTGGAGAG 1200
DB 6042 TCAGACACATTAAGGAGGCTCATGACTCTGCAAAATTTGGTGAACGTTTGGAGAG 6101
QY 1201 GAGTCTTCTTTCAGGATAGAGAGATCTTTCGAGGACACATCTCTGCTGATGAT 1259
DB 6102 GAGTCTTCTTTCAGGATAGAGAGATCTTTCGAGGACACATCTCTGCTGATGATGAT 6161
QY 1260 ACATGAGTTTCTCATGATTTGGACCTTATCAAGCTAATGGAAATAGATTTGGCGGG 1319
DB 6162 ACATGAGTTTCTCATGATTTGGACCTTATCAAGCTAATGGAAATAGATTTGGCGGG 6220
QY 1320 GTTGGTCCAGTTTGAAGGCGGTTTATGAGATCTCTGCTGATCAACCGATA 1379
DB 6221 GTTGGTCCAGTTTGAAGGCGGTTTATGAGATCTCTGCTGATCAACCGATA 6280
QY 1380 TGAAGAAATCAAGCGATGTGCTTGAAGAAATCTTCAAGCTTCAAGTCCGATC 1439
DB 6281 TGAAGAAATCAAGCGATGTGCTTGAAGAAATCTTCAAGCTTCAAGTCCGATC 6340
QY 1440 TGAAGTGTTAACGGTGTGTCTGTAGCCAGGCGATATGCTTCAAGTCAAGGCGAT 1499
DB 6341 TGAAGTGTTAACGGTGTGTCTGTAGCCAGGCGATATGCTTCAAGTCAAGGCGAT 6400
QY 1500 TCAGAGAAAAAGACAAAAATTAAGATTAAGGCTTAAGACGGGATATCTGAATTTATGA 1559
DB 6401 TCAGAGAAAAAGACAAAAATTAAGATTAAGGCTTAAGACGGGATATCTGAATTTATGA 6460
QY 1560 TAAGTGTGTGTACATCTGCACTGCAAAATATCCAACTCAGGCAATGCTGCAATGCA 1619
DB 6461 TAAGTGTGTGTACATCTGCACTGCAAAATATCCAACTCAGGCAATGCTGCAATGCA 6520
QY 1620 TACCAATATTTTCAAGGACCACTGAACCAAGCGGTTGATTAACAGCCATATGACAGATC 1679
DB 6521 TACCAATATTTTCAAGGACCACTGAACCAAGCGGTTGATTAACAGCCATATGACAGATC 6580
QY 1680 GTCAAAACCTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739
DB 6581 GTCAAAACCTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 6640
QY 1740 TGTCTCATGAGACGGGATGCAAAAGCAAGTGAATGCTCTGACATATGATGATGATGAT 1799
DB 6641 TGTCTCATGAGACGGGATGCAAAAGCAAGTGAATGCTCTGACATATGATGATGATGAT 6700

QY 1800 GA 1801
Db 6701 GA 6702

RESULT 6
TIMPS2
LOCUS
DEFINITION T1 plasmid (A. tumefaciens octopline strain) tms1 gene, complete
cde
TIMPS2 2591 bp DNA linear BCT 19-MAR-1996
T1 plasmid (A. tumefaciens octopline strain) tms1 gene, complete
K02554
K02554.1 GI:154747
2 of 2
plasmid T1
plasmid T1
plasmid
1 (bases 1 to 2591)
Klee, H., Montoya, A., Horodyski, F., Lichtenstein, C., Garfinkel, D.,
Fuller, S., Flores, C., Peachon, J., Nester, E. and Gordon, M.
Nucleotide sequence of the tms gene of the pTiA6NC octopline T1
plasmid: two gene products involved in plant tumorigenesis
Proc. Natl. Acad. Sci. U.S.A. 81 (6), 1728-1732 (1984)
84170374
MEDLINE
PUBMED
6584906
Original source text: Plasmid T1 (clone: pTiA6NC.).
The tms1 protein has a region which showed significant homology to
the adenine binding region of p-hydroxybenzoate hydroxylase from
P. fluorescens. It seems that the tms1 protein binds adenine
either as substrate or cofactor.
The initiation and termination sites of transcription have not yet
been identified [1].

FEATURES
source
Location/Qualifiers
1..2591
/organism="Plasmid T1"
/mol_type="genomic DNA"
/specific_host="octopline"
/db_xref="taxon:2512"
/clone="pTiA6NC."
/plasmid="Plasmid T1"
160..2427
/gene="tms1"
160..2427
/gene="tms1"
/codon_start=1
/transl_table=11
/protein_id="AA02550.1"
/db_xref="GI:154750"
/translation="MSASPLLDNQCCHLPTKMYDLTWVDKADLDRVSDAPLERENS
RGRITQISTEGSAGLAKRLADGRPEISAGKVAIVASVYIGKILGRILSKPM
ARATVGLVADLAPFCMDREBAOLIOALFLSGKCAPIDLSFPAVLSIKTAGPPT
LMPVLVNGTKMKTGFTITLLEGAVPPDVAAYGRNMLKKSASSPTIIDLYRRPF
DOCSDSGRIGFPEVDKPKVAIVAGISLVANELLHAGVDVITYESDVGGFL
WSHAFRAPSVAEMGMRPPAPFCLFFELERYGLSSMPPFNGTVDYLYVQVQ
YMMKAGQLPKLPHRYVNGWRAFLKGFYRDIIVLSPVAITOLKSGDVRMAHDSQ
ITINRGRSESSGIERIFLGTHPGEETSFPDMDFLTMGIGSGGFGFVPSGFI
ETIRLVINGEENORMCEGISLPRIASBVNGVSQRIHVORAIQKXKTXIK
IRKSGISELYDKVYVTSGLANTQLRCLCTDNRIPAPVNOAVNDSHMTGSKPLM
TERKFWLDHILPSCVLMDGIAKAVYCLDYPQDPNGGLVLSYTWEDSHKLLAVPD
KKEKRLCLRDALISRSPPAFQHLFPACADYDQVITQDMVLTDRNAGAFPLNRGDF
YSEELFQALDNTANDTGVYLAGCSGFTGMEVGAQIOTACNAVCAIITHNGGILAKN
PLEHSMKRYVYRTRN"

BASE COUNT 654 a 570 c 656 g 711 t
ORIGIN About 250 bp after segment 1.

Query Match 97.7%; Score 1760.4; DB 1; Length 2591;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1789; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY 1 ATGTGAGCTTCACTCTCTGTTGATTAACGAGTGCATATCTCCCAACCAAAATGCTGAT 60
Db 160 ATGTGAGCTTCACTCTCTGTTGATTAACGAGTGCATATCTCCCAACCAAAATGCTGAT 219

QY 61 CTGACATGTCGATTAAGCGCGATGATTGACCGAGGGTTCCGATGCTTCTTAGAA 120
|||
Db 220 CTGACATGTCGATTAAGCGCGATGATTGACCGAGGGTTCCGATGCTTCTTAGAA 279
|||
QY 121 CGAAGCTTCTTAGGGAAGAGATTACTCAAGCTCCACGAGTGCAGCGCTGGTTA 180
|||
Db 280 CGAAGCTTCTTAGGGAAGAGATTACTCAAGCTCCACGAGTGCAGCGCTGGTTA 339
|||
QY 181 GCTTGAAAAGGTGGCCGATGTCGCTTCCCGAGATCTCAGCTGTGGAAGTAGCA 240
|||
Db 340 GCTTGAAAAGGTGGCCGATGTCGCTTCCCGAGATCTCAGCTGTGGAAGTAGCA 399
|||
QY 241 GTTCTCTCGCTTATATCTATATTTGGCAAGAAATCTGGGGCGGATCTTGAATGAAA 300
|||
Db 400 GTTCTCTCGCTTATATCTATATTTGGCAAGAAATCTGGGGCGGATCTTGAATGAAA 459
|||
QY 301 CTTTGGGCGGGCAACAGTGAAGTGTCTGCTTGCATGCACTTGGACACCATTTGCATG 360
|||
Db 460 CTTTGGGCGGGCAACAGTGAAGTGTCTGCTTGCATGCACTTGGACACCATTTGCATG 519
|||
QY 361 GATTCTCGAAGCAACAATAATCCAAAGCCGTTTGTGCGGGTAAAGATGCA 420
|||
Db 520 GATTCTCGAAGCAACAATAATCCAAAGCCGTTTGTGCGGGTAAAGATGCA 579
|||
QY 421 CCGATTGATCTTATGATCTTCTGCGCAATTCCTAAGACTGCGGCTTTCGAAAC 480
|||
Db 580 CCGATTGATCTTATGATCTTCTGCGCAATTCCTAAGACTGCGGCTTTCGAAAC 639
|||
QY 481 CTGCGAATGCGCTGTGACGAATGGCAAGTAAGTAATGCGTTACCGGTTTACATPACC 540
|||
Db 640 CTGCGAATGCGCTGTGACGAATGGCAAGTAAGTAATGCGTTACCGGTTTACATPACC 699
|||
QY 541 CTTGAAAGGGGCGGTGCATTTGACATGATGATGATGATGATGATGATGATGATGATGAT 600
|||
Db 700 CTTGAAAGGGGCGGTGCATTTGACATGATGATGATGATGATGATGATGATGATGATGAT 759
|||
QY 601 TCGGCAAGTTCTCTTTCCAACAATCGACTGCTACGACGAGACCGTTTTCGACCA 660
|||
Db 760 TCGGCAAGTTCTCTTTCCAACAATCGACTGCTACGACGAGACCGTTTTCGACCA 819
|||
QY 661 TGTTCGATATGTCAGCGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGAAAGTGGCG 720
|||
Db 820 TGTTCGATATGTCAGCGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGAAAGTGGCG 879
|||
QY 721 GTCATTTGGGCGTGGCATTTCCGGACTCGTGGTGGCAAGCAAGTGCCTTATGCTGGGGTA 780
|||
Db 880 GTCATTTGGGCGTGGCATTTCCGGACTCGTGGTGGCAAGCAAGTGCCTTATGCTGGGGTA 939
|||
QY 781 GACGATGTTACATATATATGAAAGCAAGTGAATCGTGTGAGGCAAGCTTTGTCATATGCT 840
|||
Db 940 GACGATGTTACATATATATGAAAGCAAGTGAATCGTGTGAGGCAAGCTTTGTCATATGCT 999
|||
QY 841 TTAGGAGAGCTCTTGTGTGTGTGGCGCAAAATGGGGCCATGTCGATTTCTCTGCTGCA 900
|||
Db 1000 TTAGGAGAGCTCTTGTGTGTGTGGCGCAAAATGGGGCCATGTCGATTTCTCTGCTGCA 1059
|||
QY 901 TTCTGCTGTTTTTCTTCTCGAGAGGTACGGCTGTCTGATGAGGCGGTTCCCAAT 960
|||
Db 1060 TTCTGCTGTTTTTCTTCTCGAGAGGTACGGCTGTCTGATGAGGCGGTTCCCAAT 1119
|||
QY 961 CCGGCAAGTGCACACTTACTTGTCTACCAAGGGTCCCAATCATGTGGAAGCGGG 1020
|||
Db 1120 CCGGCAAGTGCACACTTACTTGTCTACCAAGGGTCCCAATCATGTGGAAGCGGG 1179
|||
QY 1021 CAGCTCCACCGAAGCTTTCATCGCGTTTACACAGTGTGGCGTGTCTTGAAGGAC 1080
|||
Db 1180 CAGCTCCACCGAAGCTTTCATCGCGTTTACACAGTGTGGCGTGTCTTGAAGGAC 1239
|||
QY 1081 GGTTCCTAGAGGAATATGTTGTTGGCTTGGCTGCTGCTATTACTAGGCGCTTGAAA 1140
|||
Db 1240 GGTTCCTAGAGGAATATGTTGTTGGCTTGGCTGCTGCTATTACTAGGCGCTTGAAA 1299
|||

QY	1141	TCAGAGCAATTAGGTGGGCTCATGACCTCCGCGAAATTTGGCTGAAACCGTTTGGGAGG	1200
Db	1300	TCAGAGCAATTAGGTGGGCTCATGACCTCCGCGAAATTTGGCTGAAACCGTTTGGGAGG	1359
QY	1201	GAGTCCTCTCTTCAGGGATAGAGAGATCTTTCTGGGCAACATCCCTCGTGGT-AA	1259
Db	1360	GAGTCCTCTCTTCAGGGATAGAGAGATCTTTCTGGGCAACATCCCTCGTGGTGA	1419
QY	1260	ACATGAGATTTTCTCATGATTTGGGACCTATTCAGCTAATGGGAATAGATCTGGCGGG	1319
Db	1420	ACATGAGATTTTCTCATGATTTGGGACCTATTCAGCTAATGGGAATAGATCTGGC-GG	1478
QY	1320	GTTTGTCCTCACTTTTGAAGCGGGTTTATTGAGATCTCCGCTTGTCATCAACGATA	1379
Db	1479	GTTTGTCCTCACTTTTGAAGCGGGTTTATTGAGATCTCCGCTTGTCATCAACGATA	1538
QY	1380	TGAAGAAATAGCGGATGTGCCCTGAAGGAATCTCAGAACTTCACGTGGGATCGCATC	1439
Db	1539	TGAAGAAATAGCGGATGTGCCCTGAAGGAATCTCAGAACTTCACGTGGGATCGCATC	1598
QY	1440	TGAAGTGTTAAACGGTGTGTCTGTAGCCAGCGCATATGCATGTTCAAGTCAGGCGAT	1499
Db	1599	TGAAGTGTTAAACGGTGTGTGTAGCCAGCGCATATGCATGTTCAAGTCAGGCGAT	1658
QY	1500	TCAGAGGAAAGAACAATAATTAAGATTAAGGCTTTAAGCGGGATATCTGAACTTTTGA	1559
Db	1659	TCAGAGGAAAGAACAATAATTAAGATTAAGGCTTTAAGCGGGATATCTGAACTTTTGA	1718
QY	1560	TAAAGTGTGTGCATCTGATCGCAATTCACACTCAGGCAATGCTCGAATCGGA	1619
Db	1719	TAAAGTGTGTGCATCTGATCGCAATTCACACTCAGGCAATGCTCGAATCGGA	1778
QY	1620	TACCAATATTTTTAGGCAAGGCAAGTGAACCAAGCGGTGATTAAGGCCATATGACAGATC	1679
Db	1779	TACCAATATTTTTAGGCAAGGCAAGTGAACCAAGCGGTGATTAAGGCCATATGACAGATC	1838
QY	1680	GTCAAAACCTCTTCGTATGACTGAACGAATAATTCGTGTAGACCATATCCTCCCGTCTG	1739
Db	1839	GTCAAAACCTCTTCGTATGACTGAACGAATAATTCGTGTAGACCATATCCTCCCGTCTG	1898
QY	1740	TGTCCTCATGAGCGGATCGCAAAAGCAGTATTTGCTGGACATATGATCGACAGATCC	1799
Db	1899	TGTCCTCATGAGCGGATCGCAAAAGCAGTATTTGCTGGACATATGATCGACAGATCC	1958
QY	1800	GA 1801	
Db	1959	GA 1960	

	RESULT 7
	AB016260
LOCUS	AB016260 206479 bp DNA circular BCT 12-JUN-2003
DEFINITION	Agrobacterium tumefaciens plasmid pTi-SAKURA, complete sequence.
ACCESSION	AB016260 AB007145 AB007188 AB006857 AB006858
VERSION	AB016260.1 GI:6498173
KEYWORDS	.
SOURCE	Agrobacterium tumefaciens (Rhizobium radiobacter)
ORGANISM	Agrobacterium tumefaciens
	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
	Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
REFERENCE	1
AUTHORS	Hattori,Y., Suzuki,K., Uraji,M., Ohra,N., Karoh,A. and Yoshida,K.
TITLE	Genome structure of pTi-SAKURA (t) : strategy for DNA sequencing of a Japanese cherry-Ti plasmid
JOURNAL	Nucleic Acids Symp. Ser. 37, 159-160 (1997)
MEDLINE	98247140
PUBMED	9586048
REFERENCE	2
AUTHORS	Hattori,Y., Suzuki,K., Ohra,N., Uraji,M., Karoh,A. and Yoshida,K.
TITLE	Genome structure of pTi-SAKURA (Y) : Complete nucleotide sequence of plasmid pTi-SAKURA's vir region in Agrobacterium tumefaciens
JOURNAL	Nucleic Acids Symp. Ser. 39, 265-266 (1998)
REFERENCE	3

FEATURES	source
gene	gene
CDS	CDS
COMMENT	gri:2982699, gri:2982703.
REFERENCE	Location/Qualifiers
AUTHORS	1. .206479
JOURNAL	/organism="Agrobacterium tumefaciens"
MEDLINE	/mol_type="genomic DNA"
PUBMED	/strain="MAF301001"
TITLE	/db_xref="taxon:358"
JOURNAL	/plasmid="pTi-SAKURA"
MEDLINE	/note="collection site: Japan (from cherry tree)"
PUBMED	join(205959. .206479,1. .1435)
AUTHORS	/gene="tiorf195"
TITLE	join(205959. .206479,1. .1435)
JOURNAL	/note="hyub gene homolog"
MEDLINE	/gene="tiorf195"
PUBMED	/note="hyub gene homolog"
AUTHORS	/codon_start=1
TITLE	/transl_table=11
JOURNAL	/protein_id="BAAB7820.1"
MEDLINE	/db_xref="GI:6498368"
PUBMED	/translation="MDAVRYTAAVNNRNNMIVVEASAIYRTATTEFKIVODYOCAIA TAGCEMAYPMLSGVAVFVSPPLKPTLDIAIGENLRKGDIIINDFPATGLVTHLMD VTLTLPFPEDEKILAIQMAFVHASDIDGAVPGRISPAFEPQEGRLVPMKIEGVG LNIAIKSIDODNSRIPELMDGIQAMISGLKSDRVSLETCERDYSQSEVEGNOJYN YATKRAVIRTRIDGVYTESDYLBEIGHDQLVYFSTVMTVKQSELEVDFTGDPOLA AAANLVGATTHPYIIIOCLYAYILTVDPDIPRNSGILRIAHGAPGVTAANVAPSG GSSAAATRAIVDIVILGCLNOLPEGLAAAGGSGVIVASDPRIKGRVNVNVTGD GGGARGVVDGLDGESEYVRSORSPVEVIEIEVLVWMLRIALVPSRGRVAFSGAL EIEEMTNSRAVITVNNIRFVAPVAFPGKEGILGKAVNVNNGRDERDSVGKISVLE LGGDILIRITSSTGARGFELERDVAAIRAEIENGMLSPERADVYAVVFPDRSKDVE AATAEHRERKANKSDPNCFCIVROKODVNSQOTRKLASRLTTEONIRSOLOVDR HHRLHLKGERVDKQKDDQVIAEEAGRL"
gene	gene
CDS	CDS
COMMENT	gri:2982699, gri:2982703.
REFERENCE	Location/Qualifiers
AUTHORS	1. .206479
JOURNAL	/organism="Agrobacterium tumefaciens"
MEDLINE	/mol_type="genomic DNA"
PUBMED	/strain="MAF301001"
TITLE	/db_xref="taxon:358"
JOURNAL	/plasmid="pTi-SAKURA"
MEDLINE	/note="collection site: Japan (from cherry tree)"
PUBMED	join(205959. .206479,1. .1435)
AUTHORS	/gene="tiorf195"
TITLE	join(205959. .206479,1. .1435)
JOURNAL	/note="hyub gene homolog"
MEDLINE	/gene="tiorf195"
PUBMED	/note="hyub gene homolog"
AUTHORS	/codon_start=1
TITLE	/transl_table=11
JOURNAL	/protein_id="BAAB7820.1"
MEDLINE	/db_xref="GI:6498368"
PUBMED	/translation="MSYQIEIGSRFCFVVAABELHRRARIRLNMSPPLTSOIKRLE ELMGVDFSTRRDVTRGASFFPHARDILKIDETALIMAKULEAGGSGPGEELIT GATSPITSLPLILSRFORPRTPLAIKEVDSIELGLIERGDPHVLGMRPNSGN LVFPHILSRFAVAVROSALRKALRLSPAGHGVALKKFEELSCFETVYEQLSG


```
AGLIENDIIVSDTNTALATVAGVGITFLPEWVSAMAPDGVVTRVVDMSVEISMAY
GMPDPCVPGILPEVEFATLAONLISARK"
complement (2594. .3319)
/gene="t1orf2"
complement (2594. .3319)
/gene="t1orf2"
/functions="involved in nopaline transport"
/note="Probable nocm gene"
/codon_start=1
/transl_table=11
/protein_id="BAA87627.1"
/db_xref="GI:6498175"
/translation="MDIQLIESFPKLLAAPTLLTAFISLIGFVSVPAVMRLS
KNHVSILAYGVYIRSTPLVOMFLIYVSAQFQVLSGVLSMSFREPFCALIA
LAINTAATVEIRIGGISVLSQIIEAARVAGSTPLQFERIVPIAIRALPAYGNE
VMLIKSTSLASTITIVETGLAKOITISATYSPVEVITVAGAIYLFITFVVSLLMLA
EMWLNPMARAVGGTAPKAAETH"
complement (3321. .4031)
/gene="t1orf3"
complement (3321. .4031)
/gene="t1orf3"
/functions="involved in nopaline transport"
/note="Probable nocm gene"
/codon_start=1
/transl_table=11
/protein_id="BAA87628.1"
/db_xref="GI:6498176"
/translation="MDITLQMGDAGCDELARGAMNTVVVAACSVPRGIIIFGSLPA
AKISRFWSLRLDGVYTVVVRGVELIITLVFGGGLTKRTANGLEGIEITPEPI
FVIGVLCISVSAGVATEVIRAVALPQOIEAKSIGMPWLRRLRVILPQARPA
LPGIGVMOFTLQDTSLSVGLVEIWRITAMAGSGTKQPFYITAFVIFLLSSVS
NRGLKAEKXANGVRSGO"
complement (4106. .4957)
/gene="t1orf4"
complement (4106. .4957)
/gene="t1orf4"
/functions="involved in nopaline transport
(nopaline-binding protein)"
/note="Probable nocm gene"
/codon_start=1
/transl_table=11
/protein_id="BAA87629.1"
/db_xref="GI:6498177"
/translation="MKFENLALAAVTVGLAAPTQAKDVYSITATEGSYAPNF
KDAGKLGIDLDLNDLCRMIIECKFEVQAWDGIIPSLTAGRYDAIMAAKIQPAR
EKYIARSPLYLTPMTPLTTADSEPLTQVAIENLPIDNTTPQKALDKPTIFBEV
KFGVQAGTSHAFMKQMPBVQISTYDTIDNVMDLAKRIDSLASVSLKEP/ITDIP
DNDLKMFGPRMTGGLFGKGVGIRKEDLALFDKAIIDAIADGTQKLSQOMFG
YDASPKQ"
complement (5022. .5795)
/gene="t1orf5"
complement (5022. .5795)
/gene="t1orf5"
/functions="involved in nopaline transport"
/note="Probable nocm gene"
/codon_start=1
/transl_table=11
/protein_id="BAA87630.1"
/db_xref="GI:6498178"
/translation="MDATQPTLVAEDVHKNFTLEVLKGISLTANKGDVVSIISSGGS
GKSTFLACNMFLETPNKGRIAVQGEVVVTTDAAGRLIGVDRKKIEMRWQLGVPFS
FNIMGHMTVLQNWEGFLHYKQPKGEVRADMFIDKVIANKHAAVPSQLSGGQOQ
RVSIRARLAMPBAILFDEPTSLDPELVSEVVKVTRKLABEGRTWVVVTHENGFPAD
VSSKVLFLKQGLTQCTPOEVFONPISPCRAFLSSVL"
complement (6016. .6951)
/gene="t1orf6"
complement (6016. .6951)
/gene="t1orf6"
/functions="transcriptional regulator for nopaline
catabolism"
/note="Probable nocm gene"
/codon_start=1
/transl_table=11
```

```
/protein_id="BAA87631.1"
/db_xref="GI:6498179"
/translation="MRPCVFMKEKAMIOSROLEAPRAVMLTGMTSAMLVRITOPAI
SRIRLDEEIGISLPERTCNRLPRREAGLIEKVSRRFNGIOHIDKAAELKSKHM
GSLRVACYAPALPSFMSGVITQTIADRPVSVYLDVPSQTVLELHSDYDGLSIL
AGDYPGLTTEPVPSFRAVCLLPGRHLDEKETHAHDLGESLILCISPVSLMQYDA
ALDSCEVHCNRIEISLALNCLDLVSRGSGVGVDPEDTADYSANVIOIRSPDPVVPY
HEAIVLTPSPPRPLVSEFRALDLKALPVETI"
complement (7308. .8624)
/gene="t1orf7"
complement (7308. .8624)
/gene="t1orf7"
/functions="involved in conjugation"
/note="Probable trb1 gene"
/codon_start=1
/transl_table=11
/protein_id="BAA87632.1"
/db_xref="GI:6498180"
/translation="MIQSLKGTQAGEBALSGMRINRLPIIIVVLVAFVIGVIE
GLTSRLYFRGDEGPDTSNGNPASTPADQIKGVSDGIIIGEPOLVTLQPTPVEIOE
KRAVNPFTQPGRREARPEAOLESQVWRARLQREBOYLREOROMARLQANDA
AYDSPIAVDPDKFAGRGADTAMASATTGVTATTSASDLVYALRLGIGGQNDP
NGSKKEDFPMADLKEGLYLPNRVVPQGSYIEIKRSVYPATLITGINDLPRIHQ
VSQNVYDSATGRHLIPQGTLLFGRYDSVVSFQSKRVLVVWTDIIFPNSTLDIGGA
GIDAOQSGSRDKVRKMLQTFGSAILLVIGGIDMAVSESTLADDTASDAARN
FAETFGVADRTIORMDVQPTLEIRPGYFNVLDQDIVFPQTVR"
complement (8637. .9113)
/gene="t1orf8"
complement (8637. .9113)
/gene="t1orf8"
/functions="involved in conjugation"
/note="Probable trbH gene"
```

```
Query Match      89.1%; Score 1605.2; DB 1; Length 206479;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1692; Conservative 0; Mismatches 108; Indels 2; Gaps 2;

QY      1 ATGTCACTTCACTCTCTCTGATTAACCAAGTGCATCTCCCAACCAAAATGTGTGAT 60
Db      188587 ATGTCACTTCACTCTCTCTGATTAACCAAGTGCATCTCTCAACCAAAATGTGTGAT 188646

QY      61 CTGCAATAGTGTGATTAAGCGGATGAATTGAGCCGAGGTTCCGATGCTTCTTAGAA 120
Db      188647 CTGATATAGTGTGATTAAGCGTGTGATGAATTGAGCCGAGGTTCCGATGCTTCTTAGAA 188706

QY      121 CGAAGACTTCTGAGGAAGAGAGATTACTCAAGCTCCACGAGGCGAGCGTGGGTTA 180
Db      188707 CGTGAAGCTTCTGAGGAAGAGAGATTACTCAAGCTCCGAGGCGAGCGTGGGTTA 188766

QY      181 GCTTGCAAAAGCGTGCAGATGCTCGCTTCCCGAGATCTCAAGCTGTGAAAAGTAGCA 240
Db      188767 GCTTGCAAAAGCGTGCAGATGCTCGCTTCCCGAGATCTCAAGCTGTGAAAAGTAGCA 188826

QY      241 GTTCTCTCCGCTTATATCTATATGTGGCAAGAAATCTGGGCGGATCTTGAATGAAA 300
Db      188827 GCCCTCTCCGCTTATATCTATATGTGGCAAGAAATCTGGGCGGATCTTGAATGAAA 188886

QY      301 CTTTGGCGCGGCAACAGAGTGTCTCGTTGSCATGAGCTTGAACACCATTTGCATG 360
Db      188887 CTTTGGCGCGGCAACAGAGTGTCTCGTTGSCATGAGCTTGAACACCATTTGCATG 188946

QY      361 GATTTCTCGAAGACAACTAATCCAAGCCCTTTTCTGAGCGGTAAAGATGTGCA 420
Db      188947 GATTTCTCGAAGACAACTAATCCAAGCCCTTTTCTGAGCGGTAAAGATGTGCA 189006

QY      421 CCGATTGATCTTATGATTTGCGGCAATTTCAATCTTAAAGCTCGGCGTTTGAACC 480
Db      189007 TCCAGGATCTTATGATTTGCGGCAATTTCAATCTTAAAGCTCGGCGTTTGAACC 189066

QY      481 CTCGCAATGCGGCTGACGAAATGCGACAGATGAATGCGTTACCGGTTTACCATAAC 540
Db      189067 CTCGCAATGCGGCTGACGAAATGCGACAGATGAATGCGGTTTACCATAAC 189126
```

QY 541 CTTGAAGGGCCGTCATTTGACATGTAAGCTTATGTCGAAACCTGATGCTGAAGGCT 600
Db 189127 CTTGAAGGGCCGTCATTTGACATGTAAGCTTATGTCGAAACCTGATGCTGAAGGCT 189186
QY 601 TCGGCAGGTTCTTTCCAAACAATGACATGCTGCTCAACATGCGAGACCGTTTGTGACCA 660
Db 189187 TCGGCAGGTTCTTTCCAAACAATGACATGCTGCTCAACATGCTGCTTGTGACCA 189246
QY 661 TGTTCGATGTGAGAGGATCGGCTCTTCCGAGGATGATGCTTAAGCCGGAAGTGGCG 720
Db 189247 TGTTCGATGTGAGAGGATCGGCTCTTCCGAGGATGATGCTTAAGCCGGAAGTGGCG 189306
QY 721 GTCATTGCGCTGCAATTTCCGACCTGCTGAGCAACGAACTGCTTCAATGCTGAGGTA 780
Db 189307 ATCATTTGGCCCTGCAATTTCCGACCTGCTGAGCAACGAACTGCTTCAATGCTGAG 189366
QY 781 GACGATTTACATATATGAAAGCAAGTATGCTGTTGAGGCAAGCTTTGCTCATGCT 840
Db 189367 GACGATTTACATATATGAAAGCAAGTATGCTGTTGAGGCAAGCTTTGCTCATGCT 189426
QY 841 TTCAAGGACGCTCCTAGTGTGTCGCGGAATGGGGGCGATGCGATTTCTCTGCTGCA 900
Db 189427 TTCAAGGACGCTCCTAGTGTGTCGCGGAATGGGGGCGATGCGATTTCTCTGCTGCA 189486
QY 901 TTCCTGTTGTTTCTTCTCTGACGCTTACGCGCTGCTTCTGATGAGGCGTTCCCAAT 960
Db 189487 TTCCTGTTGTTTCTTCTCTGACGCTTACGCGCTGCTTCTGATGAGGCGTTCCCAAT 189546
QY 961 CCCGGCAGTGCACACTTACTTGTGCTACCAAGCGCTGCATATCATGTGGAAGCCGGG 1020
Db 189547 CCCGGCAGTGCACACTTACTTGTGCTACCAAGCGCTGCATATCATGTGGAAGCCGGG 189606
QY 1021 CAGTGCACGCAAGCTGCTCCATCGCGGTTTACACGCTTGGTGGTCTTGAAGGAC 1080
Db 189607 CAGTGCACGCAAGCTGCTCCATCGCGGTTTACACGCTTGGTGGTCTTGAAGGAC 189666
QY 1081 GGTTCATGAGCAGATATTTGTTGTTGCTGCTGCTGCTTATTAATCAAGCCTTGA 1140
Db 189667 GGTTCATGAGCAGATATTTGTTGTTGCTGCTGCTGCTTATTAATCAAGCCTTGA 189726
QY 1141 TCAGGACACATTAAGTGGGCTCATGACTCTGCGCAAAATTTGGCTGAACCGTTTCG 1200
Db 189727 TCAGGACACATTAAGTGGGCTCATGACTCTGCGCAAAATTTGGCTGAACCGTTTCG 189786
QY 1201 GAGTCTCTCTCTCAGGAGATGAGAGATCTTCTGCGGACACATCTCTGCTGCT-AA 1259
Db 189787 GAGTCTCTCTCTCAGGAGATGAGAGATCTTCTGCGGACACATCTCTGCTGCTGAA 189846
QY 1260 ACATGAGTTTCTCTCATGATGGGACCTTATTCAGGCTTAATGGGAATAGGATCGGCGG 1319
Db 189847 ACATGAGTTTCTCTCATGATGGGACCTTATTCAGGCTTAATGGGAATAGGATCGGCGG 189905
QY 1320 GTTTGCTCAGTTTGAAGCGGGTTTATTAAGATCTTCGCTGCTGCTCAACGGA 1379
Db 189906 GTTTGCTCAGTTTGAAGCGGGTTTATTAAGATCTTCGCTGCTGCTCAACGGA 189965
QY 1380 TGAAGAAATACAGCGGATGTCCTGGAAGAAATCTCAAACTTCAAGCTGAGTCGATC 1439
Db 189966 TGAAGAAATACAGCGGATGTCCTGGAAGAAATCTCAAACTTCAAGCTGAGTCGATC 190025
QY 1440 TGAAGTGTGTAACGATGCTGCTGAGCAGCGCATATGTCATGTTCAATGCGGGGAT 1499
Db 190026 TGAAGTGTGTAACGATGCTGCTGAGCAGCGCATATGTCATGTTCAATGCGGGGAT 190085
QY 1500 TCAGAAAGAAAGCAAAATAAAGATTAAGCTTAAGAGCGGATATCTGAACCTTATGA 1559
Db 190086 TCAGAAAGAAAGCAAAATAAAGATTAAGCTTAAGAGCGGATATCTGAACCTTATGA 190145
QY 1560 TAAGGTGTGTGATGATGCTGAGCTCGCAATATTCAACTAGGCTTGTGACATGCGA 1619
Db 190146 TAAGGTGTGTGATGATGCTGAGCTCGCAATATTCAACTAGGCTTGTGACATGCGA 190205
QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAAGCATATGACAGATC 1679

Db 190206 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAAGCATATGACAGATC 190265
QY 1680 GTCAAACTCTTCTGATGATGACGAAATATTTGTTGATGACATATCTCCGCTTG 1739
Db 190266 GTCAAACTCTTCTGATGATGACGAAATATTTGTTGATGACATATCTCCGCTTG 190325
QY 1740 TGTCTCATGAGAGGATGCAAAACAGTGTATGCTGACTATGAGTCCGAGATCC 1799
Db 190326 TGTCTCATGAGAGGATGCAAAACAGTGTATGCTGACTATGAGCAGGATCC 190385
QY 1800 GA 1801
Db 190386 GA 190387
RESULT 8
AE007927
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1. 9931
/organism="Agrobacterium tumefaciens str. C58 (Cereon)"
/mol_type="genomic DNA"
/strain="C58"
/db_xref="taxon:181661"
/plasmid="71"
240. .905
/gene="AGR_pti_bx83"
240. .905
/gene="AGR_pti_bx83"
/note="putative indole-3-lactate synthase"
/codon_start=1
/transl_table=1
/product="AGR_pti_bx83p"
/protein_id="AAK90667.1"
/db_xref="GI:15163472"
/translation="MYHSRPIFNIIIDSSNIQDRLEIKVLRHTEIAYRSFAOEDLIPA
OSWMSIINTDVPIDPAIDEVVKRCEVACLPAGIPLNITLDSLTYYVCSFQAM
RYAHKRFDGVSDEGVISTVPAEGITKTRSMHNVCNTSNETHDIDAYIAF
LPTSLQNPFSHMKIGCDSFLAPSRVDPCEVILAVGKALPFDNGLKKNPRGGQWLP
PSC"
complement(1410. .2813)
/gene="AGR_pti_48"
complement(1410. .2813)
/gene="AGR_pti_48"
/note="(AJ237586) IAH"
/codon_start=1
/transl_table=1
/product="AGR_pti_48p"
/protein_id="AAK90668.1"
/db_xref="GI:15163472"
/translation="WVPTSLAQTIERLRKDYSCUELVETLIARQAKPLNALAT

DMGLRRSAKKIDRHNGNAGLGCIGIPLCFKANIATGIPFSAATPALINHLPKIPSRV
AELPFSAGALPGASGNHLSFGITSNUNNATGVARNWPNPILRGSSSGGAAVAASR
LMIGIGTGTGASVRLPALCGVGRPTIARPRRIIPVSTRETTAGIINACVADV
IIDQVTSRSAKISPMPLKRLIGLPTTFYDDLDADFAFAETIRRLIANGVADV
EADI PHLEINLSASLPIALYERPPALKKYLDVFGVSSDVIKGRSPDVANIVSA
QIOGHQISNDEVELAROSPRLQATYRNFRLYQDAILFPAPLAAKAGCESSVI
HNSMNTFKIYRVNVDPSNAGLPGSLPACLTIPRPLVGMEDIDGLAGSDHRLAIG
AALEKXINPSPFPDAPN"
3151..5418
/gene="AGR_pt1_49"
3151..5418
/gene="AGR_pt1_49"
/note="AF126446" tryptophan monooxygenase"
/codon_start=1
/transl_table=11
/product="AGR_pt1_49p"
/protein_id="AAK90969.1"
/db_xref="GI:15163473"
/translation="MSASALLDNOCDFSTKMTVDLIMVDKADELDRRVSDFSEREAS
RGRITQISGECSAGLACKRLADGRPEISAGKVALASVYVKEILLRIIESEPM
AKRVSGLVADLAPFCMDFSEKQLOTLTLGSKRCSADLSHPALVSIKTRBSRT
LQMPYEKGTTRKVTGFTLTLEAVPPDMVAYGRNMLKASAGSPFTIDILVYRSF
DQSDSGRIGFPEEDVPEKRVAIIGAGISGLVVASLHAGVDVITYEASDVGGL
WSHAFKQAPSVVEMGMRPAPASCIFFLERYGLSSMFPNPGTVDNLYQGIK
YMKAGQOPKLFHRVYSGWRAPLKDGFHGDIVLASPAITQALSKSGDIRRAHDSMO
TWINRGRESSESAIERIFLGTHPGEETSPDHMDLFLMGIISGSGRGPVSESGFI
EITRLVINGEENORMCSBGISELPRIATQVNVGVSQRITHVOYRAIEKRTKIK
IRKSGISELYDKVVVTSGLANTQRLRLCTDITIIPAPVQAVDNSHMTGSKLFL
TERKFWLIDHILPSCVLMIDGIAKAVYCLDYPQDNGGLVLISETMEDSHKLLAED
KKERFCLIRDAISRSFPAPQHLVPACADYQDQNDHMLIDENAGACFPLNKGEDF
YSEELFOALDMTNDTNGVYLACGSCSFTGMEVAGIQTACNAVCALIHNGGILLADN
PLEHSMKRYVYRNRN"
6089..6838
/gene="AGR_pt1_50"
6089..6838
/gene="AGR_pt1_50"
/note="Probable tmr gene"
/codon_start=1
/transl_table=11
/product="AGR_pt1_50p"
/protein_id="AAK90970.1"
/db_xref="GI:15163474"
/translation="MYSRYRKLMIDLRLIFGPTCTGKTSYPAVALAQOTGLPVLSLDR
OCCPOLSTGSRPTVEELKGTSTLYLDRPLVIGITAKAQAHERIMEEYVNVYAHGL
ILBEGSISLCKMAQSSYMSADPRMHIIRLELADBEETPMVAARVYQMLRPAAGSI
IOELVDLWKEPRLPIKEIDGIRYALPASQNTISDMLQDADMEDKLINGIAOE
YLIHARQEQKFRVNAAYDGFGEHFGMY"
7354..7914
/gene="AGR_pt1_51"
7354..7914
/gene="AGR_pt1_51"
/note="Probable 6a gene"
/codon_start=1
/transl_table=11
/product="AGR_pt1_51p"
/protein_id="AAK90971.1"
/db_xref="GI:15163475"
/translation="WTRQGMVLPCLSHKDDQLODELSELSKAYRKFFQTDLTKSD
IIDPSGEFLVYLDEENYRIQROSMVLSNAPDGLATTLEPVSSTYTFQVQVQLQA
FSGDGRINSRNEHSSYFLTIQASNEFGVAGVAVNTFGQSDIKWRMPDASOPLD
YHILAVGSAFLPEBALDVESDCAV"
complement (8397..9020)
/gene="AGR_pt1_bx9p"
complement (8397..9020)
/gene="AGR_pt1_bx9p"
/note="tmi"
/codon_start=1
/transl_table=11
/product="AGR_pt1_bx9p"
/protein_id="AAK90972.1"
/db_xref="GI:15163476"
/translation="MTVANQVRDLTLILRTGEMOSRLEQARTFDGALLPETYFQPS
AIRLGEFDDEYILRLQELVYVYLREDIARQCALRRHLPSNNSNGINAVATAPMLMDA

RLNRYMORCDOGLVHYQPHNTQFLAIPNSCFVRFGTIVINNENYGFYARG
NTEBEDDDDEMDDEDEFGAETRSQGNLINVYIALGSCSHLSA"
BASE COUNT 2639 a 2177 c 2364 g 2751 t
ORIGIN
Query Match 89.0%; Score 1603.6; Db 1; Length 9931;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1691; Conservative 0; Mismatches 109; Indels 2; Gaps 2;
QY 1 ATGTCAAGCTTCACTCTCTCTGATTAACAGTGGATCATCTCCCAACCAAAATGGTGGAT 60
DB 3151 ATGTCAAGCTTCACTCTCTCTGATTAACAGTGGATCATCTCTCAAAATGGTGGAT 3210
QY 61 CTGACAAATGCGATTAAGGCGATGAATGSAACCGAGGGTTCCGATGCTCTTAA 120
DB 3211 CTGATATAGTCATTAAGGCGATGAATGSAACCGAGGGTTCCGATGCTCTCA 3270
QY 121 CGAAGAGCTTCTAGGGGAAGAGATTACTCAAGCTCAACGAGTGCAGCGCTGGGTTA 180
DB 3271 CGTGAAGCTTCTAGGGGAAGAGATTACTCAAGCTCAACGAGTGCAGCGCTGGGTTA 3330
QY 181 GCTTGCAAAAGCTGGCCGATGTCCTTCCCGAATCTCAGCTGTGAAAGGTAGCA 240
DB 3331 GCTTGCAAAAGCTGGCCGATGTCCTTCCCGAATCTCAGCTGTGAAAGGTAGCA 3390
QY 241 GTTCTCTCGGCTATATCTATATTGGCAAGAAATCTGGGGCGGATCTTGAATCGAAA 300
DB 3391 GCCCTCTCGGCTTACATCTATGTTGGCAAGAAATCTGGGGCGGATCTTGAATCGAAA 3450
QY 301 CCTTGGCGCGGCAACAGTAACTGTCCTGTTCCATCGACTTGGACACCATTTGCA 360
DB 3451 CCTTGGCGCGGCAACAGTAACTGTCCTGTTCCATCGACTTGGACACCATTTGCA 3510
QY 361 GATTTCTCGAAGCAACATTAACCAAGCCCTGTTTCTGACGGGTAAAGATGCA 420
DB 3511 GATTTCTCGAAGCAACATTAACCAAGCCCTGTTTCTGACGGGTAAAGATGCA 3570
QY 421 CCGATTGATCTTATGATTCATTTGCGGCAATTCATCTTAAGACTCGGCTTTGCA 480
DB 3571 TCCAGGATCTTATGATTCATTTGCGGCAATTCATCTTAAGACTCGGCTTTGCA 3630
QY 481 CTGCAATGCGCGTGTAGCGAATGCGAATGCGAATGCGGTTTACCATPACC 540
DB 3631 CTGCAATGCGCGTGTAGCGAATGCGAATGCGAATGCGGTTTACCATPACC 3690
QY 541 CTTGAAAGGGCGCTGCCATTTTGCATGATGCTTAATGCTGAAACCTGAGCTGAAAGGT 600
DB 3691 CTTGAAAGGGCGCTGCCATTTTGCATGATGCTTAATGCTGAAACCTGAGCTGAAAGGT 3750
QY 601 TCGGCAAGGTTCTTTCCAAACAATCGACTGCTCAAGCTGACGACCGTTTTTGA 660
DB 3751 TCGGCAAGGTTCTTTCCAAACAATCGACTGCTCAAGCTGACGACCGTTTTTGA 3810
QY 661 TGTTCGATAGTGAAGGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGAAATGGCG 720
DB 3811 TGTTCGATAGTGAAGGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGAAATGGCG 3870
QY 721 GTTATTGGCGCTTGGCATTTTCCGATCTGCTGTGCGAAGCAACTGCTTACGCTGGGTA 780
DB 3871 GTTATTGGCGCTTGGCATTTTCCGATCTGCTGTGCGAAGCAACTGCTTACGCTGGGTA 3930
QY 781 GAGCATGTTCAATATATGAAGCAAGTATCGTGTGAGAGCAAGCTTTGGTCAATGCT 840
DB 3931 GAGCATGTTCAATATATGAAGCAAGTATCGTGTGAGAGCAAGCTTTGGTCAATGCT 3990
QY 841 TTCAGGAGCGCTCTAGTGTCTGCGCAATGGGCGCATGCGATTTCTCTGCTGCA 900
DB 3991 TTCAGGAGCGCTCTAGTGTCTGCGCAATGGGCGCATGCGATTTCTCTGCTGCA 4050
QY 901 TTCTGCTGTTTTTTTCTTCTGAGCGCTTACGCGCTGCTTCAATGAGCGCTTCCCAAT 960
DB 4051 TTCTGCTGTTTTTTTCTTCTGAGCGCTTACGCGCTGCTTCAATGAGCGCTTCCCAAT 4110

QY 961 CCCGGCAGCTGACACTTACTGTTCTACCAAGCGCTCCAAATACATGTGAAAGCCGGG 1020
 DB 4111 CCGGCGACAGCTGACACTTACTGTTCTACCAAGCGCTCCAAATACATGTGAAAGCCGGG 4170
 QY 1021 CACTGCGCAGCAAGCTGTTCCATCGGTTTACACGGTTGGCGGTCTTTGAAGGAC 1080
 DB 4171 CACACGACCGCAAGCTGTTCCATCGGTTTACACGGTTGGCGGTCTTTGAAGGAC 4230
 QY 1081 GATTTTCATGAGCAGATATGTTGGCTCGCTGCTGCTTACTGACGGCCCTTGA 1140
 DB 4231 GATTTTCATGAGCAGATATGTTGGCTCGCTGCTGCTTACTGACGGCCCTTGA 4290
 QY 1141 TCAGACACATTAAGTGGGCTCATGATCTCTGGCAATTTGGCTGAACCTTTGGGAGG 1200
 DB 4291 TCAGACACATTAAGTGGGCTCATGATCTCTGGCAATTTGGCTGAACCTTTGGGAGG 4350
 QY 1201 GATTCCTTCTTTCAGGAGATAGAGAGATCTTTCTGGGCAACATCTCTGCTGCT-AA 1259
 DB 4351 GATTCCTTCTTTCAGGAGATAGAGAGATCTTTCTGGGCAACATCTCTGCTGCTGAA 4410
 QY 1260 ACATGAGTTTCTCTCATGATTTGGGACCTATTCAAGCTAATGGGAATAGGATCTGGCGG 1319
 DB 4411 ACATGAGTTTCTCTCATGATTTGGGACCTATTCAAGCTAATGGGAATAGGATCTGGC-GG 4469
 QY 1320 GTTTGGTCCAGTTTGTGAAGCGGTTTATTAGATCTCTCGCTGGTCTCATCAACGGATA 1379
 DB 4470 GTTTGGTCCAGTTTGTGAAGCGGTTTATTAGATCTCTCGCTGGTCTCATCAACGGATA 4529
 QY 1380 TGAAGAAATCAGCGGATGTCCTCTGAAGGAATCTCAGAACTTCCAGCTCGATTCGCATC 1439
 DB 4530 TGAAGAAATCAGCGGATGTCCTCTGAAGGAATCTCAGAACTTCCAGCTCGATTCGCATC 4589
 QY 1440 TGAAGGTTTAAAGGTGTCTGTGAGCCGCGCATATGCGATGTTCAATCGCGCGAT 1499
 DB 4590 TGAAGGTTTAAAGGTGTCTGTGAGCCGCGCATATGCGATGTTCAATCGCGCGAT 4649
 QY 1500 TCAGAGGAAAGCAAAATAAAGATTAAGCTTAAAGCGGATATCTGAACCTTTATGA 1559
 DB 4650 TCAGAGGAAAGCAAAATAAAGATTAAGCTTAAAGCGGATATCTGAACCTTTATGA 4709
 QY 1560 TAAAGTGTGTGTCAATCTGAGCTCGCAATATCAACTCAGGCAATTCCTGACATGCGA 1619
 DB 4710 TAAAGTGTGTGTCAATCTGAGCTCGCAATATCAACTCAGGCAATTCCTGACATGCGA 4769
 QY 1620 TACCAATATTTTACGACCACTGAAACCAAGCGGTTGAATAACCATATGAACGATC 1679
 DB 4770 TACCAATATTTTACGACCACTGAAACCAAGCGGTTGAATAACCATATGAACGATC 4829
 QY 1680 GTCAAACTCTTCTGATGATGACGAAATTTCTGGTATGACCATATCTCCCGTCTTG 1739
 DB 4830 GTCAAACTCTTCTGATGATGACGAAATTTCTGGTATGACCATATCTCCCGTCTTG 4889
 QY 1740 TGTCTCATGAGCGGATCGCAAAAGAGTGTATGCTGAGTATGAGTTCGAGAGATCC 1799
 DB 4890 TGTCTCATGAGCGGATCGCAAAAGAGTGTATGCTGAGTATGAGTTCGAGAGATCC 4949
 QY 1800 GA 1801
 DB 4950 GA 4951

RESULT 9
 AEO09419 10377 bp DNA linear BCT 20-DEC-2001
 LOCUS Agrobacterium tumefaciens str. C58 Ti plasmid, section 2 of 21 of
 DEFINITION the complete sequence.
 AEO09419 AEO08690
 ACCESSION AEO09419.1 GI:17744019
 VERSION
 KEYWORDS
 SOURCE Agrobacterium tumefaciens str. C58 (U. Washington)
 ORGANISM Agrobacterium tumefaciens str. C58 (U. Washington)
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.

REFERENCE
 AUTHORS
 1 (bases 1 to 10377)
 Wood, D.W., Seubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.,
 Zhou, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F.,
 Zhou, Y., Boye Sr., D., Chapman, P., Clendenning, J., Deatherage, G.,
 Gillet, W., Grant, C., Guenther, D., Kueyavin, T., Levy, R., Li, M.,
 McClelland, E., Palmieri, A., Raymond, C., Rouse, G.,
 Saenphimmachak, C., Wu, Z., Gordon, D., Sisen, J.A., Paulsen, I.,
 Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M.,
 Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
 Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V.,
 and Nester, E.W.
 The genome of the natural genetic engineer Agrobacterium
 tumefaciens C58
 Science 294 (5550), 2317-2323 (2001)
 MEDLINE
 PUBMED
 11743193
 REFERENCE
 2 (bases 1 to 10377)
 Wood, D.W., Seubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.,
 Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F.,
 Zhou, Y., Boye Sr., D., Chapman, P., Clendenning, J., Deatherage, G.,
 Gillet, W., Grant, C., Guenther, D., Kueyavin, T., Levy, R., Li, M.,
 McClelland, E., Palmieri, A., Raymond, C., Rouse, G.,
 Saenphimmachak, C., Wu, Z., Gordon, D., Sisen, J.A., Paulsen, I.,
 Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M.,
 Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
 Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V.,
 and Nester, E.W.
 Direct Submission
 Submitted (27-SEP-2001) Department of Microbiology, University of
 Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA
 98195-7242, USA
 FEATURES
 source
 Location/Qualifiers
 1..10377
 /organism="Agrobacterium tumefaciens str. C58 (U.
 Washington)"
 /mol_type="genomic DNA"
 /strain="C58"
 /db_xref="taxon:180835"
 /plasmid="Ti"
 319..513
 /gene="ipt"
 /note="synonym: Atu6006"
 319..513
 /gene="ipt"
 /note="identified by sequence similarity; putative; ORF
 located using Blastx/Glimmer"
 /codon_start=1
 /transl_table=1
 /product="isopentenyladenosine phosphotransferase"
 /protein_id="AAL46242.1"
 /db_xref="GI:17744020"
 /translation="MDLVLFEGPTGTGTAATLSEHTGIPVALDRIQCPOLSTG
 SGRPMAEYTAPOSIGOKRC"
 complement(380..1414)
 /gene="mas1"
 /note="synonym: Atu6007"
 complement(380..1414)
 /gene="mas1"
 /note="identified by sequence similarity; putative; ORF
 located using Blastx/Glimmer"
 /codon_start=1
 /transl_table=1
 /product="agrocinnopine synthase reductase"
 /protein_id="AAL46243.1"
 /db_xref="GI:17744021"
 /translation="MHEGLGPRKQFVENYPPCEKTELFSARVAGALFTHARSNILEV
 SHGVLVAVATLIGVELTAMORCNACVLPKFDGSSMNVHTPDSGVILVTGASGGIDK
 AIDNLMRGRFLSGARSVELTEATPFGQNSLHAPRFEADLTOMEMVSTAMIKF
 GRIDGLVNNAGDGRVALABGIDYQLERQMKINAGAPURKIMLCPQOLASGSGRIV
 NITISGQRIANSPFVGNMTKIALSGLTAKTQHVGVESGVRIMEICPGVATKMSAMT
 DEIEDDGRIQEPHIELVROAIERDNNAFVPEKIEVLICIQPFLAVRSRVSQSIYVGP
 KRPCCQCAH"
 complement(1622..2878)

gene

gene
/gene="acex"
/note="synonym: Atu6008"
complement (1622. .2878)
/gene="acex"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="agrocinnopine synthase"
/protein_id="AAL46244.1"
/db_xref="GI:17744022"
/translation="MEGTVPAPWTLHMDLEPNTPEPVLARSHSPILLERLEEDLPLQ
VIRHGWFLNGKRIQOECTASLLAALQGGRLNSELVCLTSDNVAIVSHDLNTWVS
EKLGDKLFNEHSSKIKDVPVIREVNGIIOKYEITIDHILPTEIFSKVLANS
ATIFLDGRNEAVHIVAMLSHREPHORVVVLFPYEPYHGAFLDAVANAOPASMR
KSIAMPALEPRLCLARLROVTEPVDLYLAGKAMPDSMTMODMRVYAAHVSG
VTNLLGVQVVDKVLAFDSQNAVRLAYTLKSDMTIRAKRPHLKRAVTRCDDPAL
LDSEGEFSIDIKTGRARHETDERKHIRWRKGTGNSATYADWVISDRPEDEMAIW
EMRNOGIDREVSHSLPDLNLETSK"
3636. .4157
/gene="genes5"
/note="synonym: Atu6009"
3636. .4157
/gene="genes5"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="gene 5 protein"
/protein_id="AAL46245.1"
/db_xref="GI:17744023"
/translation="MNSIINTDVPIDPAIDEVKRFCEVACLPGAPAGIPLNIIILNSL
TLYVCSQAMRKXKAHRYFDVSGDEGVISTVPYAGIKETMRSHNNVCONTSRE
THOLDATIAFLPTSLQNPSPSHMKIGCDSTLAFSRVDFCEVLIANGKALFHNGLKK
NPRGGQWLPSC"
complement (4662. .6065)
/gene="tms2"
/note="synonym: Atu6010"
complement (4662. .6065)
/gene="tms2"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="indole acetamide hydrolase"
/protein_id="AAL46246.1"
/db_xref="GI:17744024"
/translation="WVPITSLAQTLERLRRKDYSCLELVETLIARCOAKPLNALLAT
DMQGLRSAKRIDRHGNAGLGCIPLCFANATATGIFPSAATPALINHLPIPSRV
AERLFSAGALPGASGNHLSFGTSNNVATGAVRNPMPSLPGSGSGVAAVADR
LMIGGIGTDGASVRLPALCGVGPRLTARVPRDIIVSPTRDTAGIIAACVADV
IILDVYSGRSAKISMPKLGRLGPTTYFDLDDADVAPAAETTRILANGVPTV
EADIPHLBELNCSALPILAYEPHALUKLDDVGVSVSDVIKGRSPDVANIVSA
QIDGHOISNDEVELARQSFRLQATYRNRYQLDAILFPAPLAAKAIGESSVT
HNSGMNTFKIYRNVDPSSNAGLPGSLPACLPRLPVGMEIDGLAGSDHRLAIG
AALEKAINPSPDAFN"
6403. .8670
/gene="tms1"
/note="synonym: Atu6011"
6403. .8670
/gene="tms1"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="tryptophan 2-monooxygenase"
/protein_id="AAL46247.1"
/db_xref="GI:17744025"
/translation="MSASALINDNOCDFSTKMDLIWMDKADELDRVSDAERRAS
RGRRTIISGECSAGLACKRLADRPPEIFISAGEVYALSIYVKEILRIIBSEBW
ARARVSLVAILDALPFCMDPSEADQLQTLTSLGKRCASSDLHFVAISISKARST
LQMPYEKCTTKRTVGTFTLLEAVFPDMVAYGNLMLKASGDSFPTIDLLYRSFP
DQSDSGRIGTFPREDVKKPVATIGAISGLVVASLHAGVSDVTIYEASDRVGKL

MSHAFKADPSVVAEMGAMRPPAASCLFFELERYGLSSNRPFPPNPGTVDNLYVQGLR
YMKAGQCPKLFHRVYSGMRAFLKDGFEHGIIVASPAVIAQALKSGDIRAHDSMO
TWNPRGQSPFSAIERIFLGTGPPRSGEIDLEFKLNGISGGGAPVSESGFI
ETLRIVINGEENORMSGISLPRRIITOVNGVSVORIRHVORAIEXEKTIK
ILKSGISLIVKRVVTTSGLANIQLRHCTTCOTTFIRAVNVAVNSHWTSKXPL
TERKEMDLHILPSCVMDGIARVCLDTEPDPNKGLVLSYEDSHKILAVPD
KKEPFLRLDAISREPAFOHLVPACADYDNNVQHDMLTEBNAGAKLRNGEDF
VSEELFQALDMTNDTGVYLAGCSGFTGWVEGAIQTCNNAVCALIHNCGILAKDN
PLEHSWKRYNRRN"
9368. .10090
/gene="tpc"
/note="synonym: Atu6012"
9368. .10090
/gene="tpc"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="isopentenyl transferase"
/protein_id="AAL46248.1"
/db_xref="GI:17744026"
/translation="MDLRILFGPTGKTSTAVALAONGPLVLSLDRVOCCPOLSTG
SGRPYEBLAKTSRLYLDPRVLYKGIILAKOKAHERLMGVYNYEARAGLIEGGSIL
LCKMAOSYSWSADFRWHIIRHBLADETFMNAKARVKOMLPAGSLTIQELVDLAK
EHLRLILKEIDGYRYAMLPASQNTTSDMLQLDADMDKLIHGAOBYLIHAROE
OKFRVNAAYADGFEHGFPGMY"
BASE COUNT 2745 a 2376 c 2438 g 2818 t
ORIGIN
Query Match 89.0%; Score 1603.6; DB 1; Length 10377;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1691; Conservative 0; Mismatches 109; Indels 2; Gaps 2;
QY 1 ATGTGAGCTTCACCTCTCTGATTAACGAGTGATCATCTCCCAACAAATGSGAT 60
|||||
DB 6403 ATGTGAGCTTCACCTCTCTGATTAACGAGTGATCATCTCTACCAAAATGSGAT 6462
61 CTGACATGATGTCATTAAGCGGATGAATGAGCCGAGGATTCGATGCTTTAGAA 120
|||||
DB 6463 CTGATTAATGTCATTAAGCGGATGAATGAGCCGAGGATTCGATGCTTTAGAA 6522
121 CGAAGACTTCTTAGGGAAGAGAGATTACTCAACGACGATGACGCTGGGTTA 180
|||||
DB 6523 CTGGAAGCTTCTTAGGGAAGAGAGATTACTCAATCTCGGCGAGATGACGCTGGGTTA 6582
181 GCTTGAAGAGGCTGCGGATGCTGCTCCCGAGATCTCAGCTGAGGAAGTAGCA 240
|||||
DB 6583 GCTTGAAGAGGCTGCGGATGCTGCTCCCGAGATCTCAGCTGAGGAAGTAGCA 6642
241 GTTCTGCGCTTATATCTATATTTGCAAGAAATCTGGGCGGATCTTGAATCGAAA 300
|||||
DB 6643 GCCCTCTCCGCTTACATCTATATTTGCAAGAAATCTGGGCGGATCTTGAATCGAAA 6702
301 CTTTGGCGCGGCAACAGTAGTGTCTGCTTGCATGACATTTGGAACCATTTGCATG 360
|||||
DB 6703 CTTTGGCGCGGCAACAGTAGTGTCTGCTTGCATGACATTTGGAACCATTTGCATG 6762
361 GATTTCTCGAAGCAGACATTAATCCAGCCGCTTTTGTGCGGCTAAAGATGCA 420
|||||
DB 6763 GATTTCTCGAAGCAGACATTAATCCAGCCGCTTTTGTGCGGCTAAAGATGCA 6822
421 CGATTTGATCTTATGATTTCTGAGCATTTCAATCTTAAGACTCGGCTTTTCAACC 480
|||||
DB 6823 TCACGAGATCTTATGATTTCTGAGCATTTCAATCTTAAGACTCGGCTTTTCAACC 6882
481 CTGCAATGCGGCTTGAAGAAATGCGATGAATGCGGCTTTTCAACATTAACC 540
|||||
DB 6883 CTGCAATGCGGCTTGAAGAAATGCGATGAATGCGGCTTTTCAACATTAACC 6942
541 CTTGAAGGCGGCTTGAAGAAATGCGATGAATGCGGCTTTTCAACATTAACC 600
|||||
DB 6943 CTTGAAGGCGGCTTGAAGAAATGCGATGAATGCGGCTTTTCAACATTAACC 7002

QY 601 TCGGAGAGTCTCTTCCAAATCGACTGCTACAGCTGACAGCCGTTTTTGACCA 660
Db 7003 TCGGAGAGTCTCTTCCAAATCGACTGCTACAGCTGACAGCCGTTTTTGACCA 7062
QY 661 TGTTCGATAGTGACCGGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGAAATGGCG 720
Db 7063 TGTTCGATAGTGACCGGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGAAATGGCG 7122
QY 721 GTGATGGGCGCTGGCAATTCGAGCTGCTGCTGAGGCGAAACGATCTTCAATGCTGGGTA 780
Db 7123 ATGATGGGCGCTGGCAATTCGAGCTGCTGCTGAGGCGAAACGATCTTCAATGCTGGGTA 7182
QY 781 GACGATGTTCAATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTGTGATCATGCT 840
Db 7183 GACGATGTTCAATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTGTGATCATGCT 7242
QY 841 TTCAAGGAGCGCTCTAGTGTGTCGGCCGAAATGGGGCGATGCGATTTCTCTGCTGCA 900
Db 7243 TTCAAGGAGCGCTCTAGTGTGTCGGCCGAAATGGGGCGATGCGATTTCTCTGCTGCA 7302
QY 901 TTCTGCTGTTTTCTTCTCCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 7303 TTCTGCTGTTTTCTTCTCCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7362
QY 961 CCGGAGACAGTGCACACTTACTTGTCTACCAAGCGCTCCAAATCATGTGAAAGCGG 1020
Db 7363 CCGGAGACAGTGCACACTTACTTGTCTACCAAGCGCTCCAAATCATGTGAAAGCGG 7422
QY 1021 CAGCTGCGACCGAAGCTGTTCCATCGGTTTACACCGGTTGGCGTCTTGAAGAC 1080
Db 7423 CAGCTGCGACCGAAGCTGTTCCATCGGTTTACACCGGTTGGCGTCTTGAAGAC 7482
QY 1081 GGTTCATAGAGGAGATATGTTGTTGGCTTCCCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 7483 GGTTCATAGAGGAGATATGTTGTTGGCTTCCCTGCTGCTGCTGCTGCTGCTGCTG 7542
QY 1141 TCAGGACACATTAAGTGGGCTCATGACTCTGCGAAATTTGGCTGAAACGTTTCGGAG 1200
Db 7543 TCAGGACACATTAAGTGGGCTCATGACTCTGCGAAATTTGGCTGAAACGTTTCGGAG 7602
QY 1201 GAGTCTCTTCTTCAAGGATAGAGAGATTTCTGGGCAACATCTCTGCTGCTGCTG 1259
Db 7603 GAGTCTCTTCTTCAAGGATAGAGAGATTTCTGGGCAACATCTCTGCTGCTGCTGCTG 7662
QY 1260 ACATGAGTCTTCTCATGATGAGGCTTATTCAGACTTATCGGAATAGATCTGGCGG 1319
Db 7663 ACATGAGTCTTCTCATGATGAGGCTTATTCAGACTTATTCAGGAATAGATCTGGCG 7721
QY 1320 GTTGGTCAAGTTTGAAGCGGGTTATGAGATCTCTGCTGCTGCTGCTGCTGCTGCTG 1379
Db 7722 GTTGGTCAAGTTTGAAGCGGGTTATGAGATCTCTGCTGCTGCTGCTGCTGCTGCTG 7781
QY 1380 TGAAGAAATCAACGAGTGTGCTGAGGAGATCTGAGAACTTCAAGCTGATGCTGATC 1439
Db 7782 TGAAGAAATCAACGAGTGTGCTGAGGAGATCTGAGAACTTCAAGCTGATGATGCTG 7841
QY 1440 TGAAGTGTGTAACGCTGTGCTGAGCCGCGCATATGCGCATGTTCAAGCTGAGGCG 1499
Db 7842 TGAAGTGTGTAACGCTGTGCTGAGCCGCGCATATGCGCATGTTCAAGCTGAGGCG 7901
QY 1500 TCAAGGAAAGAAACAAAATTAAGATTAAGGCTTAAGAGGCGGATATGTAAGCTTTATGA 1559
Db 7902 TCAAGGAAAGAAACAAAATTAAGATTAAGGCTTAAGAGGCGGATATGTAAGCTTTATGA 7961
QY 1560 TAAAGTGTGTAACGCTGTGCTGAGCTGCGAAATATCAAGCTGAGCTGCTGATGCGA 1619
Db 7962 TAAAGTGTGTAACGCTGTGCTGAGCTGCGAAATATCAAGCTGATGCTGATGCGA 8021
QY 1620 TACCAATATTTTCAAGCAGCAGTGAACCAAGCGGTTGTAACAGCCATATGACAGATC 1679
Db 8022 TACCAATATTTTCAAGCAGCAGTGAACCAAGCGGTTGTAACAGCCATATGACAGATC 8081
QY 1680 GTCAAAACCTCTTCTGATGATGAAGCAAGAAATCTGGTTAGACATATCTCTCCGCTTG 1739

Db 8082 GTCAAAACCTCTTCTGATGATGAAGCAAGAAATTTTGTGTAAGCATATCTCTCCGCTCG 8141
QY 1740 TGTCTCATGAGACGGGATCGCAAGACAGTGTATTCCTGAGACTATGATCGCAGATCC 1799
Db 8142 TGTCTCATGAGACGGGATCGCAAGACAGTGTATTCCTGAGACTATGATCGCAGATCC 8201
QY 1800 GA 1801
Db 8202 GA 8203
RESULT 10
AF126446
LOCUS
DEFINITION
AF126446 5968 bp DNA linear BCT 25-FEB-2000
Agrobacterium tumefaciens plasmid pTiC58 agrocinopine synthase
(acs) gene, partial cds; and 5 protein (5), indole-3-acetamide
hydrolase (iaah), and tryptophan monooxygenase (iaam) genes,
complete cds.
ACCESSION
AF126446
VERSION
AF126446.1 GI:4836565
KEYWORDS
SOURCE
ORGANISM
Agrobacterium tumefaciens (Rhizobium radiobacter)
Agrobacterium tumefaciens
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
REFERENCE
1 (bases 1 to 5968)
Olsen, L., Salomone, J. Y., Helfer, A., Schmidt, J., Hamann, P. and De
Ruifray, P.
Sequence and functional analysis of the left-hand part of the
T-region from the nopaline-type Ti plasmid, pTiC58
Plant Mol. Biol. 41 (6), 765-776 (1999)
JOURNAL
MEDLINE
PUBMED
10737141
2 (bases 1 to 5968)
Olsen, L., De Ruifray, P. and Hamann, P.
Direct Subregion
Submitted (08-FEB-1999) Cell Biology, IBMP, rue du General Zimmer
12, Strasbourg 67804, France
Location/Qualifiers
FEATURES
source
1..5968
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/strain="C58"
/db_xref="taxon:358"
/plasmid="pTiC58"
/note="Ti plasmid; T-DNA"
complement(1..19)
/gene="acs"
complement(1..19)
/gene="acs"
/codon_start=1
/transl_table=1
/product="agrocinopine synthase"
/protein_id="AAD30486.1"
/db_xref="GI:4836566"
/translation="MFNLCG"
782..1462
/gene="5"
782..1462
/gene="5"
/function="may catalyze synthesis of indole-3-lactate"
/codon_start=1
/transl_table=1
/product="5 protein"
/protein_id="AAD30487.1"
/db_xref="GI:4836567"
/translation="MYHSRPNIIDSSNIODRRLKLVLRTEIAYRSPABDLIPA
QKSWNSITNDVPLDPAIDVYKRFCEVACLPAGIPLNITLNDLSLYVCSFQAM
RKYAKRFEDVSDGVISTVPAEGLTKETKMSMHNVCNNTSNEHTDIDATYAF
LPTSLQNSFSHMKIGCDSFLAPSRVDPCVEIIVGKALFHDNGLKKPKWMSAP
TFLKRLV"
complement(1953..3353)
gene

CDS
/gene="iaaH"
complement (1953 . .3353)
/gene="iaaH"
/function="synthesizes indole-3-acetic acid from
indole-3-acetamide"
/codon_start=1
/transl_table=11
/product="indole-3-acetamide hydrolase"
/protein_id="A03048.1"
/db_xref="GI:4836568"
/translation="MPTISLAQTLERKRYSCLELVKTLARCOAARPLNALAT
DMGLRRTAKIDRHGAVGLGCIPLCEFGNATATGIPPSAATPALINLPATPSVR
POTLPFGAPLPAGSNGMHELSFGITSNVATGAVRNPWPSLPGSSGNGVAASR
LMIGICTDGAIVRLPALCGVGFPTLARYDRDIIPVSPRTDAGIIACQVAD
IIDOVISGRSAKILPMPKLRIGLPTTFYDDLDADVAFAETTRILANGVTFV
EADIPHELENSGASLPALYERPHALKYLDPEVGFELFTFORISPDVATVAG
IDHOISNDEYELAROSFRRLDATTNTRFLYOLDIIFPTAPLAKAIGOSSYTH
NGSMWMTFKIYVRNVDPSNAGLPGLSLPCLTPDLPLVGMELIDLAGSDHRLAIGA
ALBKAINFSPDPAFN"
3690 . .5957
/gene="iaaM"
3690 . .5957
/gene="iaaM"
/function="synthesizes indole-3-acetamide from tryptophan"
/codon_start=1
/transl_table=11
/product="tryptophan monooxygenase"
/protein_id="A03048.1"
/db_xref="GI:4836569"
/translation="MSASALDNDQCHDFSTKWDILMVDKADLDRRVSDFEEREAS
RGMITQISEGCSAGLACKRLADGRPELSAGEKVAASLYIVGKEILRIESEBW
ARAVISGLVADLAPCMDFSEADQLQTLISGKSCASDLHFVAISLKYARST
LOMPYEKGTTRKVTGFTLLEAVPDMVAAYRNMLKASGSPFIDILYRSPF
DOCSDRIGGFPEPVKPKVATIGAGISGLVVASLHAGVDPTVYASDVGKL
WSARFDABSVVAMKMRFPPLASCFPLEBRYGSSNRPFPNPGTVDTNLYQGR
YMKAGQPPKLFHRVYSGWRAPLKDFHGDVILASPAVATLALSGDIRRAHDSQ
TMNRGRSEFSSAIERIPLGTHPFGGETHSFPHMDLFLMGIISGSGFGEVSEGI
EILRLVNGYEENRMCSEGISLPRIATOVNVGSORIRHVVQALAEKRTKX
IRKSGISELYDKVYVSGSLANTOLRHCLCTTIPRAPNOAVDNHMGSKFLK
TERKFMIDHLPSCVLMDGIAKAVYCDYVPOBNGKGLIISYTMEDDSHKLLAPD
KERFCLLRDAISRSFPAPFAOHLVPRACADVDQNVVQDMLTDENAGAFULNRGDF
YSESLFQALDMDNTDGVYLAGSCSFTGWEVGAIGTACNAVCAIILHNGGILADN
PLEHSMKRYVYRNRN"
BASE COUNT 1543 a 1333 c 1472 g 1620 t
ORIGIN
Query Match 89.0%; Score 1602; DB 1; Length 5968;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1690; Conservative 0; Mismatches 110; Indels 2; Gaps 2;
QY 1 ATGTCAGCTCACCTCTCTTGATTAACGAGGATCATCTCCCAACCAAAATGATGAT 60
DB 3690 ATGTCAGCTCACCTCTCTTGATTAACGAGGATCATCTCTTACCAAAATGATGAT 3749
QY 61 CTGACATGATCGATTAAGCGGATGAAATGGAACGCGAGGGTTTCCATGCTTTGAA 120
DB 3750 CTGATTAATGATCGATTAAGCGGATGAAATGGAACGCGAGGGTTTCCATGCTTTGAA 3809
QY 121 CGAGAACTTCTAGGGAAGAAGATTAATCAAGCTTCAACCGAGTCAAGCGTGGGTTA 180
DB 121 CGAGAACTTCTAGGGAAGAAGATTAATCAAGCTTCAACCGAGTCAAGCGTGGGTTA 3869
QY 181 GCTTGAAGAGCGGCGGATGATGCTTCCCGAGATCTCAGTGTGGAAGGTGCA 240
DB 3870 GCTTGAAGAGCGGCGGATGATGCTTCCCGAGATCTCAGTGTGGAAGGTGCA 3929
QY 241 GTTCTCTCCCTTATATCTATTTGCAAGAAATTTGCGGCGGATCTGATTCGAAA 300
DB 3930 GCTTCTCTCCCTTATATCTATTTGCAAGAAATTTGCGGCGGATCTGATTCGAAA 3989
QY 301 CTTTGGGCGGCGGACAGATGATGCTCTGTTGCCATTCGATTTGACACCATTTGATG 360
DB 3990 CTTTGGGCGGCGGACAGATGATGCTCTGTTGCCATTCGATTTGACACCATTTGATG 4049

QY 361 GATTTCTCGAAGCACAACATAATCAAGCCCTGTTTTGCTGAGCGGTAAAGATGTGCA 420
DB 4050 GATTTCTCGAAGCACAACATAATCAAGCCCTGTTTTGCTGAGCGGTAAAGATGTGCA 4109
QY 421 CCGATTAATCTTAATGATCTTCTGTCGTCATTTCAATCTTAAGCTGCGGCTTTGCAAC 480
DB 4110 TCCAGGATCTTAATGATCTTCTGTCGTCATTTCAATCTTAAGCTGCGGCTTTGCAAC 4169
QY 481 CTGCAATGCGCGTGTACAGAAATGGCAGATGGAATGAGTTTACCATTAAC 540
DB 4170 CTGCAATGCGCGTGTACAGAAATGGCAGATGGAATGAGTTTACCATTAAC 4229
QY 541 CTTGAAGGCGCGTGTACAGAAATGGCAGATGGAATGAGTTTACCATTAAC 600
DB 4230 CTTGAAGGCGCGTGTACAGAAATGGCAGATGGAATGAGTTTACCATTAAC 4289
QY 601 TCGGCAAGTTCTTCCCAACATCGAATGCTCTACGACTGACAGCCGTTTTTGACCAA 660
DB 4290 TCGGCAAGTTCTTCCCAACATCGAATGCTCTACGACTGACAGCCGTTTTTGACCAA 4349
QY 661 TGTTCGATATGTAAGGATCGGCTCTTTCGGAAGATGTTCTTAAGCGGAAGTGGCG 720
DB 4350 TGTTCGATATGTAAGGATCGGCTCTTTCGGAAGATGTTCTTAAGCGGAAGTGGCG 4409
QY 721 GTTATTTGGGCTGGCATTTTCCGACTCGTGTGGCAAAAGAACTGCTTCACTGCGGCTA 780
DB 4410 GTTATTTGGGCTGGCATTTTCCGACTCGTGTGGCAAAAGAACTGCTTCACTGCGGCTA 4469
QY 781 GACGATTTCAATATATGAAAGCAAGTATCGTGTGGAAGCAAGCTTTGCTCACTGCT 840
DB 4470 GACGATTTCAATATATGAAAGCAAGTATCGTGTGGAAGCAAGCTTTGCTCACTGCT 4529
QY 841 TTGAGGAGGCTCTGATGTCGAGGCGCAATGGGAGGCAATGGGATTTCTCTGCGTCA 900
DB 4530 TTGAGGAGGCTCTGATGTCGAGGCGCAATGGGAGGCAATGGGATTTCTCTGCGTCA 4589
QY 901 TTTGCTGTTTTTTCTTCTCGAGCGTTACGAGCTGCTTTCGATGAGCGGTTCCCAAT 960
DB 4590 TTTGCTGTTTTTTCTTCTCGAGCGGTTACGAGCTGCTTTCGATGAGCGGTTCCCAAT 4649
QY 961 CCGGCAAGTGCACCTTAATGCTGCTCAAGGCGTCCATATCATGTGAAAGCCGGG 1020
DB 4650 CCGGCAAGTGCACCTTAATGCTGCTCAAGGCGTCCATATCATGTGAAAGCCGGG 4709
QY 1021 CAGCTGCAACCGAGGCTTTCATCGCGTTTCAAGGTTGGCGTCTTGAAGGAC 1080
DB 4710 CAGCTGCAACCGAGGCTTTCATCGCGTTTCAAGGTTGGCGTCTTGAAGGAC 4769
QY 1081 GATTTTCATGAGCGGATATGTTGTTGCTTCCGCTGCTGCTATTACTCAGGCTTGA 1140
DB 4770 GATTTTCATGAGCGGATATGTTGTTGCTTCCGCTGCTGCTATTACTCAGGCTTGA 4829
QY 1141 TCAGAACATTAATGTTGGGCTCATGACTCTGCGAAATTTGGCTGAACGTTTGGGAGG 1200
DB 4830 TCAGAACATTAATGTTGGGCTCATGACTCTGCGAAATTTGGCTGAACGTTTGGGAGG 4889
QY 1201 GAGTCTTCTCTCAGGATAGAGAGATCTTCTGAGGCAACATCTCTGAGTGGT-AA 1259
DB 4890 GAGTCTTCTCTCAGGATAGAGAGATCTTCTGAGGCAACATCTCTGAGTGGT-AA 4949
QY 1260 ACATGAGATTTTCTCATGATTTGGGACCTATTCAGACTAATGGAATAGGATCTGGCGG 1319
DB 4950 ACATGAGATTTTCTCATGATTTGGGACCTATTCAGACTAATGGAATAGGATCTGGCGG 5008
QY 1320 GTTGTGCAAGTTTGAAGCGGTTTATGAGATCTCTCGTTGCTCATCAAGGATA 1379
DB 5009 GTTGTGCAAGTTTGAAGCGGTTTATGAGATCTCTCGTTGCTCATCAAGGATA 5068
QY 1380 TGAAGAAATCAAGGATGTCCTTGAAGAAATCTCAGAACTTCCAGCTGCAATGCGATC 1439
DB 5069 TGAAGAAATCAAGGATGTCCTTGAAGAAATCTCAGAACTTCCAGCTGCAATGCGATC 5128

QY	1440	TGAAGTGTAAAGGTTGTCTGTGAGCAGGCGATATGCCATGTTCAAGTCAGGGCGAT	1499
Db	5129	TCAAGTGTAAACGGTGTGTCTGTATGCCAGCGTATACGCCATGTTCAAGTCAGGGCGAT	5188
QY	1500	TCAGAAAGAAAGA CAAAAATTAAGATTAAGGCTTAAGACGGGATATCTGAATTTATGA	1559
Db	5189	TGAGAAAGAAAGA CAAAAATTAAGATTAAGGCTTAAGACGGGATATCTGAATTTATGA	5248
QY	1560	TAAAGTGTGTGTCACATTTGGACTTGGCAATATCCAACTCAGGCAATTGCTCTGACATCGGA	1619
Db	5249	TAAAGTGTGTGTACATCTGTGACCTCGCAAAATATCCAACTCAGGCAATTGCTCTGACATCGGA	5308
QY	1620	TACCAATATTTTTTACAGGCAAGCAGTGAACCAAGCGGTTGATTAACGCCATATGACAGATC	1679
Db	5309	TACCAACATTTTTTCTGTGACCAAGTGAACCAAGCGGTTGATTAACGCCATATGACAGGCTC	5368
QY	1680	GTCAAAACCTCTTCCTGATGATCGAACAAGAAATTCGTGTTAAACCATATCTCCGCTCTTG	1739
Db	5369	GTCAAAACCTCTTCCTGATGATCGAACAAGAAATTTGTTAAGCAATATCTCCGCTCTG	5428
QY	1740	TGTCTCTCATGGAACGGGATCGCAAAAGCAGTGTATTTGCTTGACCTATGAGTGTGACAGATCC	1799
Db	5429	TGTCTCTCATGGAACGGGATCGCAAAAGCAGTGTATCTGCTGACCTATGAGTGTGACAGATCC	5488
QY	1800	GA 1801	
Db	5489	GA 5490	

RESULT 11			
LOCUS			
ATU237588			
ATU237588	29524 bp	DNA	linear BCT 09-JUN-1999
DEFINITION	Agrobacterium tumefaciens Ti plasmid pTiC58 T-DNA region.		

KEYWORDS acrociniopine synthase;

KEYWORDS acetoacinnopine synthase; acetoacinnopine synthase, IAH; isopentenyl transferase; nopaline synthase; ornithine cyclodiaminase; protein 6a; protein 6b; torf1 gene; torf10 gene; torf11 gene; torf12 gene; torf13 gene; torf14 gene; torf15 gene; torf16 gene; torf17 gene; torf18 gene; torf19 gene; torf2 gene; torf20 gene; torf21 gene; torf22 gene; torf3 gene; torf4 gene; torf5 gene; torf6 gene; torf7 gene; torf8 gene; torf9 gene; tryptophan 2-monooxygenase.

SOURCE	ORGANISM
Agrobacterium tumefaciens (Rhizobium radiobacter)	
Agrobacterium tumefaciens	
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.	

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE
1	Gidlen, J., Terry, N., Van Montagu, M. and Vilarroel, R.	Complete nucleotide sequence of the T-DNA region of the plant tumor inducing <i>Agrobacterium tumefaciens</i> Ti plasmid pTiC58 unpublished		2 (bases 1 to 29524)

TITLE Direct Submission
SUBMITTED (24-MAR-1999) Gielen J., Genetics, VIB/UG, KU
ADDRESS Ledeganckstraat 35 Gent, 9000, BELGIUM

FEATURES	Location/Qualifiers
source	1. .29524

gene	complement (1. .575) /gene="torf1"
CDS	complement (<1. .575)

```

/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAB46625.1"
/db_xref="GI:5042180"
/translation="MAPEEDRARGLRRAFLRLRFLCPSHPDAQYVHRGDLSRVAINNF

```

gene	complement (604, .1479) /gene="torf2"
CDS	complement (604, .1479) /gene="torf2"

```

/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44626.1"
/db_xref="GI:5042181"
/translation="MVQSPSGGIGRGAGATRPICEARPTNPTGPRRLQSGVASE
LEIINNIILANGFNASNDGDELISGRRLAETLGYDTVALISDLIAGSIVSE
RGDRRLITLFTSPRYTHGVYSSNFMDPRTRHQVDAAKVALICSDIRRVYVQHA
QFKANVDQAQRGADERCEALLLEIGESPAKIVGINIAENITFTEPGGYRTAPVAAV
TFLRNADHDVALIDFKPRLPGIQLSRFPFRSVARHGIVLAGPGRILSRATATRIIV
VITDGMCSGWS"
805..1353
gene

```

gene	805.1353
/gene="torf3"	
CDS	805.1353

```

/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44627.1"
/db_xref="GI:5042182"
/translation="MIDGFHR CGNFRSLTILARICENTLLKSGNAPVVRCEMLAFQEC
RALTLVAGALSDLICGRKARYDENAGVAIVNLDQCSIDLVAPRPGVLAHVEG
EAVHGIAIGKRLKKKSEPIATLFDADTCDOIENGKGFEDVNDCLKAQDLSLEAVK
LAVRRCRIEARCDNIIVHNDIQF"
complement(1359..1886)
gene

```

gene	complement (2090. .2644) /gene="corf5"
CDS	complement (2090. .2644)

gene	complement (4009 . 5124) /gene="orf6"
CDS	complement (4009 . 5124)

```

3623. .3647
/rpe_family="r-DNA left border"
/rpe_type=DIRECT
complement(4009. .5124)
/gene="torf6"
complement(4009. .5124)
/gene="torf6"
/codon_start=1
/transl_table=1
/product="agrocinnopine synthase"
/protein_id="CAB44630.1"
/db_xref="GI:5042185"
translation="MWNLKGVCETSTLLTALGCMERRNLSELDPTLLTSDNLPVVS

```

HDENWTRISTLPDLRLRELHLSADVKNI PVIIREVSNGETIESYTVTNDII PLEPEPLLD
KVEDVNGATFFEDGRDFEAMHITAMLSRPRKROVLLFFEEYSSGAFPDAYKO
VSDPEPMRETVAMPVIFPOELPRKLIDLDIAVEDLIEGKIMWIDSLRQPMVY
AVHIAWRYKPEPOLGCDKPEIVRAADVAAYVYKNDNEVRMRPHKLATAT
RCIDFSAQLENGEKAESTINIRKGRPMRETDERKTIIRKRYGIFGNAAAIADAVISDR
SEDEMSFWEWRNKGVPFRVDHCHPLDVLIEQDKSVP"
/gene="torf7"
/db_xref="GI:5042187"
/gene="torf7"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44631.1"
/db_xref="GI:5042186"
/transl_table="MYERPVPDIIVDTWEIREREVLPKVLVAEDYCYEONTVASA
QKSWDLVANLENSPDSIDIDALNRPDAFCILGPTLEPNLILKSPMYVYSFSEB
RSCATKRFEYEGISNRGVICTPPVAGVTRDMSYMONQACVNTSGKNDLDAVIAF
LPTSLQESSYMHYTKGIVSFLAPSQDACEILCVGRALFEDRSRKEKRLNCLIKL
LNRLKLLF"
complement (7150. .8721)
/gene="torf8"
complement (7150. .8721)
/gene="torf8"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44632.1"
/db_xref="GI:5042187"
/transl_table="MENQNPVPAIERKLSWLSLHVRKILPSPELGSPYSYTEE
TVDGEADICRHSVTPLPEDBGVSHHQOICSTHOPAAHDVHAFARAPBEPAY
IPEDJADKSRFPASSENSLTDILMMPSTRNAMATOKERDDAFVILILPSIP
KEYKYLIRBSAOSIOTQEMWOLKCHPLMERGLFWIMLVIMOLAVLPIVISAAGA
EYAVLSGRPTTLOTQMLVLAIFPYPADTTLAAHLACEVGLGPIVILRELPANBQ
HEGTEPHADVAKASLCTLMQMLQNRNIRDLKISFLMRKLISTAOSDPSISY
TLRNSSTIIYSWLDTPVESNRRLIQGYKHQDHEDALRVLRVWGHADNDSPISG
KAGILVALITMAGLICSTPTDAYSGLIANAALVPLAARAAVDMNDSOSHDLLRL
FTADAGVSPSTFLVIGNSSYWLTKKSLAAGFANIHFWIAFVILFPLSLFPLMGRL
FWYVSGRLNDRSA"
/gene="torf9"
9762. .10493
/gene="torf9"
9762. .10493
/gene="torf9"
/function="phosphatase"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44633.1"
/db_xref="GI:5042188"
/transl_table="MHTSLMSYNQNMIPSGDDRPVPTIDLPANTSLEELQOGLCAV
ANYRDLPTKRIIPROOSWYRVVISEGLDEETLRATSCVPTCCSIPGAVODRILQ
HSDPIYITSLNLISQYVEKTLIRCTRKRVLASLSPRYTDIPAEILKAVLANSILQ
AYSMPLAVSCFVALIPSKKFTISMSKTTTGGAEBHFVPSRENDIIPYDIAVGKT
LFQSDSEFGLPMLSPDELQNLIST"
11195. .11890
/gene="torf10"
11195. .11890
/gene="torf10"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44634.1"
/db_xref="GI:5042189"
/transl_table="MMVYQRPVFTVLSLIRNREAKVLIGAVVYVNFVEQTLAD
AOKMWKSLVLDPDGDAVIGILTWFSYACLHGPRIGPLDTIAVNDNPIYICPRK
LEYPKRIIVSFSEISGVCSMSPDAGVTRKRYGHNLIISGSLDPALEAYVA
FLBSKFLKLPYSYVEVHNDRYVHKFPALLPGRSFHEVAVAVGLAYPAKKRPSGLCI
LRCCFTCKNTTC"
12368. .13312
/gene="torf11"
12368. .13312

/gene="torf11"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44635.1"
/db_xref="GI:5042190"
/transl_table="MMVPRPGFAPLSIESINDRMEMHFLTSAQILYHIIYRDNLLI
YOEIWKYWRGRSLAESALKNEISALYRLPFRPPTLOSIAEDEVICKDAPIYVCS
REMLSHVOSKXLCOFGRANLACTAPRYGTTEMLVCLFRLADGRVEXGTSES
DLSHRYALPTTAPFRGPRFVYIGENVQFLREGLELPHCHIIAFGRALFSEILAGS
SMDKSVGESSRLNHRASVEISRVPTMETTDNABALGTDAYVPVPPNLTPTOD

Query Match 89.0%; Score 1602; DB 1; Length 29524;

Best Local Similarity 93.8%; Pred. No. 0;

Matches 1690; Conservative 0; Mismatches 110; Indels 2; Gaps 2;

Qy	1	ATGTCAGCTTCACTCTCTCTGATTAACGATGATCATCTCCCAACCAAAATGATGAT	60
Db	19987	ATGTCAGCTTCACTCTCTCTGATTAACGATGATCATCTCTCAAAATGATGAT	20046
Qy	61	CTGACATGTCGATTAAGCGGATGATGAATGAACCGGAGGTTCCGATGCTTCTAGAA	120
Db	20047	CTGATATGTCGATTAAGCTGATGATGAATGAACCGGAGGTTCCGATGCTTCTAGAA	20106
Qy	121	CGAAGACTTCTAGGGAGAGAGATTACTCAAGCTCCACCGATGCAAGCGCTGGATT	180
Db	20107	CGTGAAGCTTCTAGGGAGAGAGATTACTCAAACTCCGAGGAGTGCAGCGCTGGATT	20166
Qy	181	GCTTGCAAAAAGCTGGCCGATGATGCTCTCCCGAATCTCACTGCTGGAAGATGCA	240
Db	20167	GCTTGCAAAAAGCTGGCCGATGATGCTCTCCCGAATCTCACTGCTGGAAGATGCA	20226
Qy	241	GTTCTCTCGCTATATATATATGCAAAATCTGGGCGGATCTTGAATGAAA	300
Db	20227	GCCCTCTCGCTATATATATGCAAAATCTGGGCGGATCTTGAATGAAA	20286
Qy	301	CCTTGGCGCGGCAACAGTGAATGCTCTGTTGCCATGCACTTGGACACATTTGCATG	360
Db	20287	CCTTGGCGCGGCAACAGTGAATGCTCTGTTGCCATGCACTTGGACATTTGATG	20346
Qy	361	GATTTCTCGGAGCACAATCTCAAGCCCTGTTTTGCTGAGCGGTAATAAGATGCA	420
Db	20347	GATTTCTCGGAGCACAATCTCTCAACCTGTTTTGCTGAGCGGTAATAAGATGCA	20406
Qy	421	CGATGATCTTATGATTTCTGATTTGCGCATTTCAATCTTAAGATGCGGCTTTGCAAC	480
Db	20407	TCCAGCATCTTATGATTTCTGATTTGCGCATTTCAATCTTAAAGATGCGGCTTTGCAAC	20466
Qy	481	CTGCCAATGCGCTGTACGAGAAATGCGATGAAATGCGTTACCGGTTTACATTAAC	540
Db	20467	CTGCCAATGCGCTGTACGAGAAATGCGATGAAATGCGTTACCGGTTTACATTAAC	20526
Qy	541	CTTGAAGGGCGCTGCTTGAATGATGCTTGAATGCTGAAACCTGATGCTGAAGGCT	600
Db	20527	CTTGAAGGGCGCTGCTTGAATGATGCTTGAATGCTGAAACCTGATGCTGAAGGCT	20586
Qy	601	TGCGCAGGTTCTTCCCAACATGCACTTGTCTAGACGCTGAGACGCTTTTGAACCA	660
Db	20587	TGCGCAGGTTCTTCCCAACATGCACTTGTCTAGACGCTGAGACGCTTTTGAACCA	20646
Qy	661	TGTTCCGATAGTGAAGGATGCGCTTTTCCGAGAGATGTTCTTAAGCCGAAATGCGC	720
Db	20647	TGTTCCGATAGTGAAGGATGCGCTTTTCCGAGAGATGTTCTTAAGCCGAAATGCGC	20706
Qy	721	GTCATGCGCGCTGCTTCCGATGCTGCTGCTGCTGCTTCTTCTGCTGCTGCTGCT	780
Db	20707	ATCATTTGCGCGCTGCTTCCGATGCTGCTGCTGCTGCTTCTTCTGCTGCTGCTGCT	20766
Qy	781	GACGATGTTACATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTTGTCATATGCT	840
Db	20767	GACGATGTTACATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTTGTCATATGCT	20826

QY 841 TTCAGGAGCCTCTAGTGTGTGTCGCGAAATGGGCGCATGCCATTTCTCTGTGCA 900
DB 20827 TTCAGGAGCCTCTAGTGTGTGTCGCGAAATGGGCGCATGCCATTTCTCTGTGCA 20886
QY 901 TTCTGCTGTTTTTCTCTCTGAGCGTTAGCGCTGTCTTTCGATGAGCGCGTTCCCAAT 960
DB 20887 TCGTGCTGTTTTTCTCTCTGAGCGTTAGCGCTGTCTTTCGATGAGCGCGTTCCCAAT 20946
QY 961 CCCGCGCAGTCGACCTTACTTGGTGTACCAAGCGCTCCAAATCATGTGGAAGCGCGG 1020
DB 20947 CCCGCGCAGTCGACCTTACTTGGTGTACCAAGCGCTCCGATTCATGTGGAAGCGCGG 21006
QY 1021 CAGCTCCACCGAAGCTGTTCCATCGCGTTTCAACCGCTTGGCGCTTTGGAAGAC 1080
DB 21007 CAGAGCCACCGAAGCTGTTCCATCGCGTTTCAACCGCTTGGCGCTTTGGAAGAC 21066
QY 1081 GGTTCCTATAGGAGATATTGTGTGGCTTCCGCTGTCCCTATTTACTCAGCGCTTGA 1140
DB 21067 GGTTCCTATAGGAGATATTGTGTGGCTTCCGCTGTCCCTATTTACTCAGCGCTTGA 21126
QY 1141 TCAGGACATTAAGTGGGCTCATGACTCGCGCAATTTGGCTGAACCGCTTGGGAGG 1200
DB 21127 TCAGGACATTAAGTGGGCTCATGACTCGCGCAATTTGGCTGAACCGCTTGGGAGG 21186
QY 1201 GAGTCTCTCTCTCAGGATAGAGAGATCTTCTGGGCAACATCTCTGGTGT -AA 1259
DB 21187 GAGTCTCTCTCTCAGGATAGAGAGATCTTCTGGGCAACATCTCTGGTGTGTGMA 21246
QY 1260 ACATGAGTTTTCTCTCATGATTGGGACCTTATCAAGCTAAATGGAAATAGATCTGGG 1319
DB 21247 ACATGAGTTTTCTCTCATGATTGGGACCTTATCAAGCTAAATGGAAATAGATCTGGG -GG 21305
QY 1320 GTTTGGTCCAGTTTTTGAAGCGGCTTATTGAGATCTCTCGCTGCTCATCAAGCGATA 1379
DB 21306 GTTTGGTCCAGTTTTTGAAGCGGCTTATTGAGATCTCTCGCTGCTCATCAAGCGATA 21365
QY 1380 TGAAGAAATCAGCGGATGTGCCCTGAAGGATCTCAGAACTTCCAGCTGGATCGCATC 1439
DB 21366 TGAAGAAATCAGCGGATGTGCCCTGAAGGATCTCAGAACTTCCAGCTGGATCGCATC 21425
QY 1440 TGAAGTGTAAACGGTGTCTGTGAGCCAGCGCATATGCGCATGTTCAAGTCAGGCGAT 1499
DB 21426 TGAAGTGTAAACGGTGTCTGTGAGCCAGCGCATATGCGCATGTTCAAGTCAGGCGAT 21485
QY 1500 TCAAGGAGAAACCAAAATTAAGATTAAGGCTTAAGAGCGGATATCTGAATTTATGA 1559
DB 21486 TCAAGGAGAAACCAAAATTAAGATTAAGGCTTAAGAGCGGATATCTGAATTTATGA 21545
QY 1560 TAAAGTGTGTGATCATCTGGAATCGCAATATTCAGGATTCGCTGACATGCGA 1619
DB 21546 TAAAGTGTGTGATCATCTGGAATCGCAATATTCAGGATTCGCTGACATGCGA 21605
QY 1620 TACCAATATTTTCAAGGACAGTGAACCAAGCGGTTGATTAAGACCATATGAAGATC 1679
DB 21606 TACCAATATTTTCAAGGACAGTGAACCAAGCGGTTGATTAAGACCATATGAAGATC 21665
QY 1680 GTCAAAACTTTTCTGTGATGATGAGCAAAATTTCTGTGTGACCATATCTCTCCGCTTG 1739
DB 21666 GTCAAAACTTTTCTGTGATGATGAGCAAAATTTTGTGTGACCATATCTCTCCGCTTG 21725
QY 1740 TGTCTCATGAGCGGATTCGCAAAAGCAGTGTATGCTGCTGAGCATATGAGTCGAGATTC 1799
DB 21726 TGTCTCATGAGCGGATTCGCAAAAGCAGTGTATGCTGCTGAGCATATGAGTCGAGATTC 21785
QY 1800 GA 1801
DB 21786 GA 21787

RESULT 12
AB025110 4660 bp DNA linear BCT 27-MAR-1999
LOCUS AB025110
DEFINITION Agrobacterium tumefaciens gene for indole acetamide hydrolase and

ACCESSION
AB025110
VERSION
AB025110.1 GI:4586311
KEYWORDS
tryptophan monoxygenase; indole acetamide hydrolase.
SOURCE
Agrobacterium tumefaciens (Rhizobium radiobacter)
ORGANISM
Agrobacterium tumefaciens
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
REFERENCE
1 (bases 1 to 4660)
Kasahara, T., Endo, S., Ebinuma, H., Sugita, K., Kawaoka, A. and
Moriwae, N.
AUTHORS
Moriwae, N.
TITLE
Agrobacterium tumefaciens plasmid PO22 indole acetamide hydrolase
(iaah), tryptophan monoxygenase (iaam) genes
JOURNAL
Published Only in Database (1999)
REFERENCE
2 (bases 1 to 4660)
Kasahara, T., Endo, S., Ebinuma, H., Sugita, K., Kawaoka, A. and
Moriwae, N.
AUTHORS
Moriwae, N.
TITLE
Direct Submission
Submitted (18-MAR-1999) Takehide Kasahara, Nippon Paper Industries
Co., Ltd. R&D Div., Wood Bio-engineering Central Research Lab.,
5-21-1, Oji, Kita-ku, Tokyo 114-0002, Japan
[E-mail: takehide@bglabode.ne.jp, Tel: 81-3-3911-5106 (ex. 292),
Fax: 81-3-3914-3350]
FEATURES
Location/Qualifiers
1..4660
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/strain="PO22"
/db_xref="taxon:358"
/complement(353..1756)
/gene="iaah"
complement(353..1756)
/gene="iaam"
/codon_start=1
/transl_table=1
/product="indole acetamide hydrolase"
/protein_id="BAA76345.1"
/db_xref="GI:4586312"
/translation="MYPIRSLAQTLERLARKYSCLELVETILARCOAKPKPNNLLAT
DWDGLRSKAKDRHNGNGLGCLPLCKKATATAGVFTSATPELNLHLPKIPSRV
ABRLFSAGALPGASNMHLSFGTISNNVATGAVNPMPSLI PGSSGGAVAASVSR
LMLGIGTGTGASVRLPALCGVGRFPLTGVPRDRILIPESPTDTAGIACQVADV
ILDDVITSGRSKISPMPLKGRILGPTTYPFDIDADADVAFAETIRLLARGVTFF
EADIDITSGRSKISPMPLKGRILGPTTYPFDIDADADVAFAETIRLLARGVTFF
QIDQOISNDEYELAROSRFRPLQATRYRFPFLYQDILFLPTAPLAAATIGESVI
HNGSMNTFKITVRNVDPSSNAGLPSLPACLTLPURLPVGMEIDGLASDRLLAIG
MALEKAINSSPPDAFN"
2096..4363
/gene="iaam"
2096..4363
/gene="iaam"
/codon_start=1
/transl_table=1
/product="tryptophan monoxygenase"
/protein_id="BAA76346.1"
/db_xref="GI:4586313"
/translation="WSASPLDNCDSHTKAVDILINDKADBLDRVSDARSERAS
RGRRTQVISGECSDAGLACRLADGRFPEISTEKVAASAVIYVGEKILGRLESBPW
AARVSGVAILDAPFCMDPSEAOQLQTLFLSGRLMLKASAGSPFIDLDYRSFF
LQMPYEKGTTRATGFTLTLEAVPDMVAAGRLMLKASAGSPFIDLDYRSFF
DCCSIOIRGTFPEDEVPRKPVAILIGAGISGLVAVSELHAGVDVATYASRVGGL
WSHARKDASVPAEMGAMPFPAPASCLPEFLRVSGLSSVAPRNCTDNTNLYOGLR
YVWAKGQDPKPLPHRYVSGMRAPLADGRHECIVIVASPVITDALKSGDIRAHDSWQ
TWLNPGRRESFSAIERIFLGTHPPEGGEWSPPHMDLFXKMGISGSGGFPVESGFI
ELIRLVINGYENQRMCSGISELPRIASOVNVGSQARLHVQVRAIIEKXKIK
ILKSGISELYDKVVVTSGLANIQLRHLCTCDTITIFRAVQADVNSHTGSSKLFL
TERKFMILHILPSVIMDGIKAVYCLDVEPDGKGLVPTTYMEDSHKLAVPD
KKEKFCILRDATSRSPAPAOHLVPAQADVDQNVQHMVLTEENAGGAPKILNRGDF
YSEELPFOALDPRNDTGYTLACSCSPFGNVEGALQTRCANVCAI IHNCGSILAKDN
PLEHSMKRYNNRN"

BASE COUNT 1174 a 1038 c 1171 g 1276 t 1 others
ORIGIN

Query Match 88.7%; Score 1597.2; DB 1; Length 4660;
 Best Local Similarity 93.6%; Pred. No. 0;
 Matches 1687; Conservative 0; Mismatches 113; Indels 2; Gaps 2;

QY 1 ATGTCACCTTCACTCTCTCTTGTATACAGTCGATCATCTCCCAACCAAAATGTGTGAT 60
 DB 2096 ATGTCACCTTCACTCTCTCTTGTATACAGTCGATCATCTCTACCAAAATGTGTGAT 2155

QY 61 CTGCAATGTGTCATTAAGCGGATGAAATGGAACCGGAGGTTTCCGATGCTTTAGAA 120
 DB 2156 CTGCAATGTGTCATTAAGCGGATGAAATGGAACCGGAGGTTTCCGATGCTTTAGAA 2215

QY 121 CGAAGACCTTCTAGGGGAGAGGATTAATCAAGCTTCAACGAGTGCAGCGCTGAGTTA 180
 DB 2216 CGTGAACCTTCTAGGGGAGAGGATTAATCAATCTCCGCGAGTGCAGCGCTGAGTTA 2275

QY 181 GCTTGGCAAAAGCTGCGGATGATGTCGCTTCCCGAATCTCACTGCTGTGAAAAGTACGA 240
 DB 2276 GCTTGGCAAAAGCTGCGGATGATGTCGCTTCCCGAATCTCACTGCTGTGAAAAGTACGA 2315

QY 241 GTTCTCTCCGCTTATCTATATATGGAAGAAATTCGCGGCGGATACCTTGAATCGAA 300
 DB 2336 GCTTCTCTCCGCTTATCTATATGGAAGAAATTCGCGGCGGATACCTTGAATCGAA 2395

QY 301 CATTGGGCGGCGGACAGATGATGATCTGCTGTCATCGAATTGACACCATTTGATG 360
 DB 2396 CATTGGGCGGCGGACAGATGATGATCTGCTGTCATCGAATTGACACCATTTGATG 2455

QY 361 GATTTCTCCGAAGACAATTAATCAAGCCCTGTTTTGCTGAGCGGTAAAGATGTGCA 420
 DB 2456 GATTTCTCCGAAGACAATTAATCAAGCCCTGTTTTGCTGAGCGGTAAAGATGTGCA 2515

QY 421 CCGATTGCTTACTGATCTGTTGCTGAGCAATTCCTAAGATGCGGCTTTGCAAC 480
 DB 2516 TCCAGCATCTTACTGATCTGTTGCTGAGCAATTCCTAAGATGCGGCTTTGCAAC 2535

QY 481 CTGCGCAATGCGCGCTGTAAGAAATGGAAGATGCAATGCAATGCGGTTTACCATTAAC 540
 DB 2576 CTGCGCAATGCGCGCTGTAAGAAATGGAAGATGCAATGCAATGCGGTTTACCATTAAC 2635

QY 541 CTTGAAGGGGCGGTCGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 2636 CTTGAAGGGGCGGTCGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2695

QY 601 TCGGACAGTTCTCTTCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 2696 TCGGACAGTTCTCTTCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2755

QY 661 TGTTCGATAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 DB 2756 TGTTCGATAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2815

QY 721 GTTATTTGGCGTGGCATTTTCCGACTGCTGTGCGAAACGAATGCTTCAATGCTGGAGTA 780
 DB 2816 ATCATTTGGCGTGGCATTTTCCGACTGCTGTGCGAAACGAATGCTTCAATGCTGGAGTA 2875

QY 781 GACGATGTTCAATATATGAAGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 840
 DB 2876 GACGATGTTCAATATATGAAGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 2935

QY 841 TTGAGGACGCTCTAGTGTGTGCGGAAATGGGCGGATGCGATTTCTCTGCTGCA 900
 DB 2936 TTGAGGACGCTCTAGTGTGTGCGGAAATGGGCGGATGCGATTTCTCTGCTGCA 2995

QY 901 TTCTGCTGTTTTTTCTCTGAGCGTTACGCGCTGCTTGCATGAGCGCTTCCCAAT 960
 DB 2996 TTCTGCTGTTTTTTCTCTGAGCGTTACGCGCTGCTTGCATGAGCGCTTCCCAAT 3055

QY 961 CCGGGCAGTCAGCACTTCTGTTGATCAAGGCGTCAATTAATGATGAAAAGCCGG 1020
 DB 3056 CCGGGCAGTCAGCACTTCTGTTGATCAAGGCGTCAATTAATGATGAAAAGCCGG 3115

QY 1021 CAGCTGCACCGAAGCTGTTTCATCGCTTTACAAACGTTGGCGTGTCTTGAAGAC 1080

DB 3116 CACACACCAACCGAAGCTGTTCCATCCGTTTACAGCGGTTGGCGTCTTGAAGGAC 3175
 QY 1081 GATTTTCATGAGGAGATATGTTGTTGGCTTGGCTGCTGCTGCTTATCTGAGGCTTGA 1140
 DB 3176 GATTTTCATGAGGAGATATGTTGTTGGCTTGGCTGCTGCTGCTTATCTGAGGCTTGA 3235

QY 1141 TCAGACACATTAAGTGGGCTCATGATCTCTGCAAAATTTGGCTGAACGTTTGGGAG 1200
 DB 3236 TCAGACACATTAAGTGGGCTCATGATCTCTGCAAAATTTGGCTGAACGTTTGGGAG 3295

QY 1201 GAGTCTTCTCTTCAAGGATTAAGAGATCTTTCTGGGACACATCTCTCTGTGT -AA 1259
 DB 3296 GAGTCTTCTCTTCAAGGATTAAGAGATCTTTCTGGGACACATCTCTCTGTGTGAA 3355

QY 1260 ACATGAGTTTCTCATGATTTGGGACCTATTCACCTATATGGGATPAGGATGCGGG 1319
 DB 3356 ACATGAGTTTCTCATGATTTGGGACCTATTCACCTATATGGGATPAGGATGCGGG 3414

QY 1320 GTTGTGTCAGTTTGTGAAGCGGTTTATTAAGATCTCTCGCTGTGTCATCAACGATA 1379
 DB 3415 GTTGTGTCAGTTTGTGAAGCGGTTTATTAAGATCTCTCGCTGTGTCATCAACGATA 3474

QY 1380 TGAAGAAATTCAGCGATGTGCTCTGAAGAAATCTCAGAACTTCCAGTCGATCCGATC 1439
 DB 3475 TGAAGAAATTCAGCGATGTGCTCTGAAGAAATCTCAGAACTTCCAGTCGATCCGATC 3534

QY 1440 TGAAGTTTAAGGTTGCTGTGAGCAAGCGATATGCAATGCAATGCAATGCAATGCAATG 1499
 DB 3535 TGAAGTTTAAGGTTGCTGTGAGCAAGCGATATGCAATGCAATGCAATGCAATGCAATG 3594

QY 1500 TCAGAGGAAAGACAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAATCTTATGA 1559
 DB 3595 TCAGAGGAAAGACAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAATCTTATGA 3654

QY 1560 TAAAGTGTGTGTCATCTGAGTCTGCAATATCAACTGAGCAATGCTGCTGATGCGA 1619
 DB 3655 TAAAGTGTGTGTCATCTGAGTCTGCAATATCAACTGAGCAATGCTGCTGATGCGA 3714

QY 1620 TACCAATTTTTTCAAGGACCAAGTGAACCAAGGTTGATTAAGCATATGACAGATC 1679
 DB 3715 TACCAATTTTTTCAAGGACCAAGTGAACCAAGGTTGATTAAGCATATGACAGATC 3774

QY 1680 GTCAAAATCTTCTCTATGATGATGAAAGAAATCTGTTAGACCATATCTCCGCTTTC 1739
 DB 3775 GTCAAAATCTTCTCTATGATGATGAAAGAAATCTGTTAGACCATATCTCCGCTTTC 3834

QY 1740 TGTCTCATGAGAGGATGCGAAAGCAAGTGTATGCTGAGTATGATGCGAGATCC 1799
 DB 3835 TGTCTCATGAGAGGATGCGAAAGCAAGTGTATGCTGAGTATGATGCGAGATCC 3894

QY 1800 GA 1801
 DB 3895 GA 3896

RESULT 13
 ATTAAAH 10200 bp DNA linear BCT 06-JUL-2002
 LOCUS A.tumefaciens Tm4 Ti plasmid DNA with TA-iaah interrupted by IS866,
 DEFINITION TA-iaah, T-1pt and T-6 gene.
 ACCESSION X56185.1 GI:39133
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 1 (bases 7402 to 8783)
 REFERENCE
 AUTHORS Bonnard, G., Thirland, B., Paulus, F., Szegedi, E. and Otten, L.
 TITLE Nucleotide sequence, evolutionary origin and biological role of a
 rearranged cytokinin gene isolated from a wide host range biotype

III Agrobacterium strain
Mol. Gen. Genet. 216 (2-3), 428-438 (1989)
MEDLINE 89313683
PUBMED 2546041
REFERENCE 2
AUTHORS Bonnard,G., Vincent,F. and Otten,L.
TITLE Sequence of Agrobacterium tumefaciens biotype III auxin genes
JOURNAL Plant Mol. Biol. 16 (4), 733-738 (1991)
MEDLINE 91329707
PUBMED 1868204
REFERENCE 3 (bases 1 to 10200)
AUTHORS Bonnard,G.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-1990) G. Bonnard, INSTITUT DE BIOLOGIE
MOLECULAIRE DES PLANTES DU CNRS, 12 RUE DU GENERAL ZIMMER, 67084
STRASBOURG CEDEX, FRANCE
FEATURES
source Location/Qualifiers
1. 10200
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/strain="Tm4"
/db_xref="taxon:358"
/plasmid="T1"
/complement(join(192..653,3369..4319))
/gene="TA-1aah"
/complement(join(192..653,3369..4319))
/gene="TA-1aah"
/codon_start=1
/transl_table=11
/product="indoleacetamide hydrolase"
/protein_id="CAA39649.1"
/db_xref="GI:39137"
/db_xref="SWISS-PROT:P25016"
/translation="MTPTSIAQTLETRRKYSCELYETLLIARCEAKALNLIAT
DWGLRRRAKIDRHGNAVGLRGIPICFPANATGI FPTTATPALINLPKIPRI
AERLFSAGALPGASGMWELSGITSNNTATGARNPWSLIPGSSGVAVAASR
LMLGIGTGTGASVRLPALCGVAGPRTGAPRDRILPVSPTPDGIAQCAVDV
VILDOVISGRPARILPLPKGRIGLPTTYFDLDADVAFAAETTRILLANGVRFV
EADPHLEDNSGASLPINAXERPHAKQYLDPPVCTVSSSDVIKEIRSDVANI
VNAQIDGHQISNAEYELAROSFRPLQAAFRNFRILRDALIFPAIPLAAKIGODS
SVTHNSMNTFKIYVARNVDPSSNAGLPGISLPLVGLTPRLPVMEIDIGLAGDHRL
ALGAALEKAINFRSPDVLN"
646..653
654..3369
/insertion_seq="IS866"
3370..3377
4485..4490
/note="Sep I junction site"
4657..6924
/gene="TA-1aam"
4657..6924
/gene="TA-1aam"
/codon_start=1
/transl_table=11
/product="tryptophane monoxygenase"
/protein_id="CAA39646.1"
/db_xref="GI:39134"
/db_xref="SWISS-PROT:P25017"
/translation="MSASSLDKQCDHFSYKIVLDIMVDAKELDRVAAAFSEBEAS
RERRISQISEGCNAGLAKRLADGRPEISAGORVAVLSAYIVGEISLRMLIEPAS
VRRIVSGLAVIDLAPSCMDISRLOLQTMNLISGKCAPDLSFPVAISSEIARERT
LQAPYEGGSKSTKSTFTVYIEAVPDMVAYRNMLKASASPTTIDLDYRFL
DKSDSGRIGTFPEVPRFVAIVAGISLVAVASELHAGVDVITYEDGVAGKL
WSHAFDAPGVAVEMGMRPPAASCLFFELERYLSMPPFNPGTVTDIVYECR
YMMKAOQPEPKLFHRYSGWHAFLKDGFLLEGDIVLASPAITALKSGDI RRAMDSQ
IYNRRGRESFAIBRI FLGTHPGEETWSFPHDMLFKLMGISGGGSPVSEGPT
EILRVINGYEENORMSGEISELPRIASQVNVGWSORISGSHVAVOVRAIIEKSKTIX
IRLKSGISLVYVVTGSLANTOLRHCLCTTIRAPVNOAVDSSHMTSGSKFEL
TERKFWFDHMLPSCVLMGFAKAVYCLIDYEPDPKNGKLVLSITYTEDSHKLLAVPD
KKEKLCILRDALISKSPFVFAQHLVPA CADVDQNVVQHDWLTDENAGREFLNRGSDP
YSEELFFQALDITNDTGVYLAGCSFTGGMVGAIAQTACNAVCAI IHNCGTILAND
PLGHPMKRYVYRNRN"
7613..8335

CDS
/gene="T-1pc"
7613..8335
/gene="T-1pc"
/codon_start=1
/transl_table=11
/product="isopenlencyl transferase"
/protein_id="CAA39647.1"
/db_xref="GI:39135"
/db_xref="SWISS-PROT:P25018"
/translation="MDLRIRFPTCTGKSTALAAQNGLPVLSIDRYQCCPOLSTG
SGRPTVEELKGTIRLYLDDRLVKGI IAEQNHREILAVYVEAHGSLISGSL
LKQMSQYWSADFRFMHII RHKLADSEETFMKAKARVKQMLCPAIGPSIQLVLYLM
EPLRLPILKEIDGYVYAMLFAQONRITPMLQLDADMEGKLIHGIADYLLIHAROE
HEFPVSAASFEGPPEGAY"
complement(9013..9636)
/gene="T-6b"
/gene="T-6b"
complement(9013..9636)
/gene="T-6b"
/codon_start=1
/transl_table=11
/protein_id="CAA39648.1"
/db_xref="GI:39136"
/db_xref="SWISS-PROT:P25019"
/translation="MTVANMQRDPRTIRLINAGELQGRLEQARTDEGALLAEIVYRHP
GATPEGDEYILTCQGLVYVYLSQIARQCLNBLPSNSNFGTVTAIPMLMDT
OTLNLTLQRCDDGGIVNYHGSRTNEFLATMLNCFVRFTDEINGASVGFYARAG
NYTERGEDNDNIGDEGAGAEIRDYQPGDLVNYPIVALSSRLSA"
BASE COUNT 2436 a 2471 c 2709 g 2584 t
ORIGIN
Query Match 81.5%; Score 1467.6; DB 1; Length 10200;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 1606; Conservative 0; Mismatches 194; Indels 2; Gaps 2;
QY 1 ATGTGAGCTTCACTCTCTTGATTAACGATGATCATCTCCAAACCAAAATGTGTAT 60
Db 4657 ATGTGAGCTTCACTCTCTTGATTAACGATGATCATCTCCAAATGTGTAT 4716
QY 61 CTGACAAATGTGATTAAGCGGATGAATGACCCGAGGTTTCCGATCCTTCTTAA 120
Db 4717 CTGATTAATGTGACAAAGCTGATGATGACCCGCGTGTCCGACCTTCTCAGAA 4776
QY 121 CGAGAAAGCTTGAAGGGAAGAGATTAATCAAGGTCACCGAGTGCAGCGTGGGTTA 180
Db 4777 CGAGAAAGCTTGAAGGGAAGAGATTAATCAAGGTCACCGAGTGCAGCGTGGGTTA 4836
QY 181 GCTTGCAAAAGGCTGCGCATGTCCTTCCCGAGATCTCAGCTGTGGAAGGTAGCA 240
Db 4837 GCTTGCAAAAGGCTGCGCATGTCCTTCCCGAGATCTCAGCGGTGAGAGGGTCGA 4896
QY 241 GTTCTCTCCGCTTATATCTATATTGGCAAGAAATTTCTGGGCGGATTAATGTAATG 300
Db 4897 GTTCTCTCCGCTTATATCTATATTGGCGAGAAATTTCTGAGGTGATCTCAACACAGA 4956
QY 301 CCTTGCGCGCGGCAACAGTGAAGTGTCTGTTGCATGATGATGAGACCAATTTGCATG 360
Db 4957 GCTTCCGTCGCAACAGTGAAGTGTCTGTTGCATGATGATGAGACCAATTTGCATG 5016
QY 361 GATTTCTCCGAAGCAACATTAATCAAGGCTGTTTCTGAGCGGTAAAGATGTGCA 420
Db 5017 GATTTCTCCGAAGCAACATTAATCAAGGCTGTTTCTGAGCGGTAAAGATGTGCA 5076
QY 421 CCGATTGATCTTAGTCAATTTGCGGCATTTCAATCTTAAGACTGCGCGGCTTGAACC 480
Db 5077 CCGAGGACCTTAGTCAATTTGCGGCATTTCAATCTTGAAGACTGCGCGGCTTGAACC 5136
QY 481 CTGCAATGCGCTGATGAGAGATGCGACGATGAATGCGTGAACGGGTTTCAATTAAC 540
Db 5137 CTGCAATGCGCTGATGAGAGATGCGACGATGAATGCGGTTTCAATTAAC 5196
QY 541 CTGGAAGGGCGGTGCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 5197 ATTGAAGAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5256

gene

QY 601 TCGGAGAGTTCCTTCCAAATCGACTGCTGACAGTGCAGACCGGTTTTTGAACAA 660
DB 5257 TCGGAGAGGCTCTTCCAAATCGACTGCTGCTAGTACACAGATGTTCTTCGACAA 5316
QY 661 TGTTCGATAGTGAAGGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 720
DB 5317 TGTTCGATAGTGAAGGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 5376
QY 721 GTGATGCGCTGCGATTTCCGAGCTGCTGAGGCAAAACGATCGCTTCACTGCGGCTA 780
DB 5377 GTGATGCGCTGCGATTTCCGAGCTGCTGAGGCAAAACGATCGCTTCACTGCGGCTA 5436
QY 781 GAGGATGTTCAATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTTGTCACATGCT 840
DB 5437 GAGGATGTTCAATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTTGTCACATGCT 5496
QY 841 TTCAGGAGCGCTCTAGTGTCTGAGCGGAAATGGGGCGATGCGATTTCTCTGCTGCA 900
DB 5497 TTCAGGAGCGCTCTAGTGTCTGAGCGGAAATGGGGCGATGCGATTTCTCTGCTGCA 5556
QY 901 TTCTGCTGTTTTTCTTCTCTGAGCGTTACGCGCTGCTTCTGATGAGCGCGTTCCCAAT 960
DB 5557 TTCTGCTGTTTTTCTTCTCTGAGCGTTACGCGCTGCTGATGAGCGCGTTCCCAAT 5616
QY 961 CCGGAGCAGTGCAGCACTTACCTTGTACCAAGCGCTCCAAATCATGTGAAAGCCGG 1020
DB 5617 CCGGAGCAGTGCAGCACTTGTGTACAGAGGTTCCGATACATGTGAAAGCCGG 5676
QY 1021 CAGTGCACCGAAGCTGTTCCATCGGTTTACACGTTGGCGTGGCTTTGAAGGAC 1080
DB 5677 CAGAGCCACCGAAGTGTTCATCGGTTTACAGCGGTTGGCGTGGCTTTGAAGGAC 5736
QY 1081 GGTGTTATGAGGAGATTTGTGTGTGCTTCCGCTGCTTATCTCAGGCTTGA 1140
DB 5737 GGTGTTATGAGGAGATTTGTGTGTGCTTCCGCTGCTTATCTCAGGCTTGA 5796
QY 1141 TCGAGCACATTAAGTGGGCTCATGATCCCGGCAAAATTTGGCTGAACGTTTCGGGAG 1200
DB 5797 TCGAGGACATTAAGTGGGCTCATGATCCCGGCAAAATTTGGCTGAACGTTTCGGGAG 5856
QY 1201 GAGTCTCTTCTTTCAGGAGTAGAGAGATCTTCTTGGGCAACATCTCTGTGTG- TAA 1259
DB 5857 GAGTCTCTTCTTTCAGGAGTAGAGAGATCTTCTTGGGCAACATCTCTGTGTG- TAA 5916
QY 1260 ACATGAGTCTTCTTCTGATGAGGACCTTATCAAGCTAATGGAAATAGATCTGGGG 1319
DB 5917 ACATGAGTCTTCTTCTGATGAGGACCTTATCAAGCTAATGGAAATAGATCTGGGG 5975
QY 1320 GTTGTGTCAGTTTGTAAAGCGGTTTATGAGATCTCCGCTTGGCTCATCAACGATA 1379
DB 5976 GTTGTGTCAGTTTGTAAAGCGGTTTATGAGATCTCCGCTTGGCTCATCAACGATA 6035
QY 1380 TGAAGAAATCAGCGGATGTGCTCTGAAGGAATCTCAAACTTCCACGTCGGATCGATC 1439
DB 6036 TGAAGAAATCAGCGGATGTGCTCTGAAGGAATCTCAAACTTCCACGTCGGATCGATC 6095
QY 1440 TGAAGTGTGTAACGCTGTCTGTGAGGACGACGATATGCCATGTTCAAGTCAGGGGAT 1499
DB 6096 TGAAGTGTGTAACGCTGTCTGTGAGGACGACGATATGCCATGTTCAAGTCAGGGGAT 6155
QY 1500 TCGAAGGAAAGCAAAATTAAGATTAAGCTTAAGAGCGGATCTGAACCTTATGA 1559
DB 6156 TCGAAGGAAAGCAAAATTAAGATTAAGCTTAAGAGCGGATCTGAACCTTATGA 6215
QY 1560 TAAAGTGTGTGATCATCTGAGCTCGCAAAATATCAACTCAGGATGCTGACATCGA 1619
DB 6216 TAAAGTGTGTGATCATCTGAGCTCGCAAAATATCAACTCAGGATGCTGACATCGA 6275
QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACGCAATATGACAGATC 1679
DB 6276 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACGCAATATGACAGATC 6335

QY 1680 GTCAAAATCTTCTCTGATGACGAAAGAAATCTGCTTAGAACCATATCTCTGCTTG 1739
DB 6336 TTCAAAATCTTCTCTGATGACGAAAGAAATTTGCTTAGAACCATATCTCTGCTTG 6395
QY 1740 TGTCTCTATGAGAGGATGCAAAAGCACTGATTTCCCTGAGTATGATCGAGATCC 1799
DB 6396 TGTCTCTATGAGAGGATGCAAAAGCACTGATTTCCCTGAGTATGATCGAGATCC 6455
QY 1800 GA 1801
DB 6456 GA 6457
RESULT 14
AYU83987
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
source
source
misc_feature
repeat_region
CDS
gene
15463 bp DNA linear BCT 02-MAR-2001
Agrobacterium vitis plasmid pTiTm4 agrocinopine synthase and (5)
genes, (laah) pseudogene, and (laah), (lpt), (6b) and (oce) genes,
complete cds.
U83987.1 GI:1814322
Agrobacterium vitis (Rhizobium vitis)
Agrobacterium vitis
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
Osten, L. and De Ruffray, P.
Sequence and functional analysis of the shoot-inducing octopline
T-DNA of Agrobacterium vitis strain C6474
Unpublished
Osten, L. and De Ruffray, P.
Direct Submission
Submitted (07-JAN-1997) Institute for Plant Molecular Biology, rue
du General Zimmer 12, Strasbourg 67084, France
Location/Qualifiers
1. 15463
/organism="Agrobacterium vitis"
/mol_type="genomic DNA"
/strain="Tm4"
/db_xref="taxon:373"
/plasmid="pTiTm4"
3857. 6588
/organism="Agrobacterium vitis"
/mol_type="genomic DNA"
/db_xref="taxon:373"
/insertion_seq="I5866"
15438. 215463
/organism="Agrobacterium vitis"
/mol_type="genomic DNA"
/db_xref="taxon:373"
/insertion_seq="I5867"
139. 15310
/note="shoot-inducing octopline T-DNA"
139. 165
/note="flanking repeat; left border of the T-DNA"
complement(373. 1629)
/codon_start=1
/transl_table=11
/product="agrocinopine synthase"
/protein_id="AAB41872.1"
/db_xref="GI:1814323"
/translation="MEGTAPAPMTLTHMDI.PENVTPPEVLARHSVPKLIBELLENLPQ
VLEHGFNDGNRIQECTASLLAALGGGRNLSLDVCLTSDNITVVSDDLTWVS
EKGDKEFNEIHSDDINNVPIIREVSNGLIDRYRETDHPLDLDELKIVSNPVS
ATIFLDGRVYEAHIVAWLSHREYVQRAVVFYREYDPGAFVDAVLPAQASMR
KSIALMPALPEBELCRLARQVSEPTVDLYLAGKAMIDSLMODMRVVAIVVSE
VSRNLLDRVNDKVLAFPSDOAAVRLAHVVEDTMIKARPHLKFAAVTRCYDFPAL
LNSGERAERSIDINTGARPHETDERKHLRMKRGTONSATIDMWISDSSEEMAIW
EMRNGIDREVSLSLPHLDVDTSE"
2235. 2918
/gene="5"


```
CDS
2235..2918
/gene="5"
/codon_start=1
/translate=1
/protein_id="AAB41873.1"
/db_xref="GI:1814324"
/translation="MDHSRPIFNVIDCSHIQDRRELEIVLLATRRAYLRPAQDLIPA
QRMSIDSIDTNPVIDEVKRPVACLPAPVGMSLNIVLNDLWVYCSFOEM
RRYACERFEGSGKGVISTVPYSGIKRAMRCOMNCHITNNEKXIDLAIYAF
LPTSLONTFSHKVIGDAFLAPSRDPFCVEIVANGKALPDNGIKKAPKAGMSMA
LTSILKRLV"
complement(3403..7530)
/gene="1aah"
/complement(join(3403..3856,6589..7530))
/gene="1aah"
/notes="interrupted by the bacterial insertion element
15866"
/pseudo
/codon_start=1
/translate=1
3857..3864
/notes="flanking repeat; left border of the 15866 insertion
element"
/rpt_type=direct
6581..6588
/notes="flanking repeat; right border of the 15866
insertion element"
/rpt_type=direct
7868..10135
/gene="1aam"
7868..10135
/gene="1aam"
/codon_start=1
/translate=1
/protein_id="AAB41874.1"
/db_xref="GI:1814325"
/translation="MSASLLDKOCDFSTKYVDLIMVDADELDERRVAAAFSEREAS
RERRISQISEGNAGLACKRLADRPETISAGORVAVLSAYIVGSEILLRILEPAS
RRTRVGLVAILDAPSCMDISRAQLOLQTNMLSGKCAPSDLSHPVAISLSETRART
LQAPYBESGLKSVTGTIVIEBVPDMVAYGRNMLKASAGSPFTIDLYRFL
DKCSDSGRIGFPEDVPRVAVIAGISGLVVASLHAGVDVTIYKGDVGGKL
MSHAFDAPGVAVEMGMRPPPAASCIFLFLERYGLSMRPFNPGVTDILYEGCR
YMKRAGQPPKLFHRVYSGWAFKDGFLDGIVLASPDATLALSGDRRAHDSMO
IYNRGRRESFSSAIBRIPLCTHPRGSETPPHMDLFLMGTISGSGRPVSESGFT
EILRLVINGEENQRMCSGISLPRIRASOVNNGSVSORIRHVOVRALEKXTIK
IRKSGISELYDKVVVYSGLANIQLRCLCTDTTIFRAPVNOAVNDSHMTGSKFLD
TERKFMFDHMLPSCVLMDFKAKAVYCLDYEPDOPNGGLVLIETWEDDSHKLALVD
KERRLCILRADISKSPFVFAOHLVPACADYDQVVOHMDLTDENAGRFKLNREGDF
YSEELFEQALDITNDTGVYLAAGSCSPGTGWEAGIQTACNAVCAIITHCGGILANDN
PLGHPMKRYVYRRNR"
10824..11546
/gene="1pc"
10824..11546
/gene="1pc"
/codon_start=1
/translate=1
/protein_id="AAB41875.1"
/db_xref="GI:1814326"
/translation="MDLRLIFGPTCTGTSTAIALAQQGTLPVLSIDRVQCCPOLSTG
SGRPVTELBELKGTRLYLDRLPVKGIIIAEOAHERLIAEYVYEAHGILLEGISL
LKCMAOSGYMSADFRWHIIRKLADETFPKAKARVYKOWLCPIGPSLIOELVYVNM
EPLRLPIKEIDGYRYAMLPAQNRITTPMDLQDLDMBKLIHIGIAQETLIHAROE
HERPVSAAAFEGFBGPPFGAY"
complement(12224..12847)
/gene="6b"
complement(12224..12847)
/gene="6b"
/codon_start=1
/gene="6b"
/translate=1
/protein_id="AAB41876.1"
/db_xref="GI:1814327"
/translation="MTVANVQVRDFTRIINAGELQGLQGLQARTDFGALLAEIYFHP
GATPEGDDEYILTGGLVYVYLSQGLARQCALNRLLPSSSSNFGTVVTAIPWLMNDT
```

```
OTLNLTLQERCDQGIIVNYHGRSTNIEFLAINTLSCFVRFGTDEINGASCYGFYARRG
NYTEBEDDDNDEIGDEAGAEIRYQGGDLVNYPIVALGSSRLSA"
complement(13779..14855)
/gene="ocs"
complement(13779..14855)
/gene="ocs"
/codon_start=1
/translate=1
/protein_id="AAB41877.1"
/db_xref="GI:1814328"
/translation="NAKVAIILAGNLALTLAGDLARLQGTSPIMAPISNRSFNDVR
CIGSLTELVGPDYGDGFQPRLEDDLGTALISGAFTPLVTPLTGOGLRELAKEFLNS
VLVALPGSATSLACKQTLVTPFAPLIVISSTSPACRVKAVLMLGVAKTEVAT
OPLSEVRKRFVLPNPNOMYOHOPASIFESNTNPPYHAGIILAAEDSTEOGILPVPK
FYRQVPQALITRVIALDERLSIUDALGESSTDPYSKNNKGNAKNRPFETYE
YALIEPTKTMHRIYLEDVYKHLIVLVEIAEIAIGOVPMKSVQOASVLENLRLT
GRGLSLNIGANNAIVVALNGV"
15287..15310
repeat_region
BASE COUNT 3873 a 3665 c 3902 g 4023 t
ORIGIN
Query Match 81.5%; Score 1467.6; DB 1; Length 15463;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 1606; Conservative 0; Mismatches 194; Indels 2; Gaps 2;
QY 1 ATGTGAGCTTCACCTTCCTTGATTAACAGTGGATCATCTCCAAACCAAAATGTGGAT 60
DB 7868 ATGTGAGCTTCACCTTCCTTGATTAACAGTGGATCATCTTCCTCAAAAATGTGGAT 7927
QY 61 CTGACAAATGATCGATTAAGCGGATGATGAGACCGGAGGATTCGATGCTTCCTAGAA 120
DB 7928 CTGATATGATGACAAAGCTGATGATGAGACCGGAGGATTCGATGCTTCCTAGAA 7987
QY 121 CGAAGACTTCTAAGGGAAGAGATTACTCAAAGCTCAACGAGTGCAGCGCTGGGTTA 180
DB 7988 CGAAGACTTCTAAGGGAAGAGATTACTCAAATTCGCGGAGTGCACACCTGGGTTA 8047
QY 181 GCTTGCAAAAGCTGGCCGATGTGCTTCCTCCGAGATCTCAGCTGTGGAAGATGACA 240
DB 8048 GCTTGCAAAAGCTGGCCGATGTGCTTCCTCCGAGATTCAGCCGCTCAGAGGGTCGCA 8107
QY 241 GTTCTTCGGCTTATATCTATATGCAAGAAATCTGSGGGGGGATCTTGTAATGAAA 300
DB 8108 GTTCTTCGGCTTATATCTATATGCAAGAAATCTGAGGGGATCTCGAACAGAA 8167
QY 301 CCTTGCGCGCGGCAACAGTGAAGTGTCTGTTGCATCGACTTGACACCAATTCGATG 360
DB 8168 GCTTGCGTGCGAACAAAGATGAGTGTCTGTTGCATCGACTTGACACCATTTGATG 8227
QY 361 GATTTCTCGAAGCAACAACTTAATCCAGCCCTGTTTCTGAGCGGTAAAGATGTGCA 420
DB 8228 GATTTCTCGAAGCTCAACTTCCAAACCAAGAAATTCCTGATGTGTAAGATGTGCA 8287
QY 421 CGATTTGATCTTAATGATTTGCGGCACTTCAATCTCAAGCTGCGGCTTGGAAAC 480
DB 8288 CGAAGCACTTAATGATTTGCGGCACTTCAATCTCAAGCTGCGGCTTGGAAAC 8347
QY 481 CTGCAATCGCGCTGTACAGAAATGACAGATGAATGCGTTTACCGGGTTTACCATACC 540
DB 8348 CTGCAATCGCGCGGTACAGAAATGACAGATGCTGTTGAAGAGGTTTACCGGGTTTACCGGTAATC 8407
QY 541 CTTGAAGGGGCGGTGCAATTTGACATGTAAGCTTATGTCGAAACCTGATGCGAAGGCT 600
DB 8408 ATTGAAGGCGGATACCATTTGACATGTAAGCTTATGTCGAAACCTGATGCGAAGGCT 8467
QY 601 TCGGCAAGTTCTCTTCCAAATGCACTGCTCAAGTGAAGCGGTTTTTGAACAA 660
DB 8468 TCGGCAAGGCTCTTTCGAAGATGCACTGCTTATGATCAAGATGTTTCTCGCAAA 8527
QY 661 TGTTCGATAGTGAACGATCGGCTTCTTTCGGAAGATGTTCTTAAGCGGAAGTGGCG 720
DB 8528 TGTTCGATAGTGGGCGGATCGGCTTCTTTCGGAAGATGTTCCAGGCGCAAAAGTGAAGG 8587
```


QY 721 GTCATTGGCGCTGGCATTTTCCGAGCTGTGTGGCAAAACGAACTGCTTCATGCTGGGGTA 780
 DB 858 GTGATTTGGGGCTGGCATTTTCCGGGCTGTGTGGCAAGCAACGCTGCTTACGCTGGGTA 8647
 QY 781 GACGATGTTAAATATATGAAAGCAAGTATGTTGTGGAGCAAGCTTTGTGCATGCT 840
 DB 8648 GACGATGTTAAATATATGAAAGCAAGTATGTTGTGGAGCAAGCTTTGTGCATGCT 8707
 QY 841 TTCAAGGAGCGCTCTAGTGTGTGGCCGAAATGGGGCCGATGGAATTTCTCTGCTGCA 900
 DB 8708 TTCAAGGAGCGCTCTAGTGTGTGGCCGAAATGGGGCCGATGGAATTTCTCTGCTGCA 8767
 QY 901 TTCTGCTTTGTTTCTCTCTCGAGCGCTTACGCGCTGCTTCAATGAGCGCTTCCCAAT 960
 DB 8768 TCGTGTGTTTCTCTCTCGAGCGCTTACGCGCTGCTTCAATGAGCGCTTCCCAAT 8827
 QY 961 CCCGGCAGCTGCACTTACTGTTGTTACCAAGCGCTCCATATCATGTGAAAGCCGG 1020
 DB 8828 CCCGGCAGCTGCACTTACTGTTGTTACCAAGCGCTCCATATCATGTGAAAGCCGG 8887
 QY 1021 CAGCTGCCAGCAAGCTGTTCCATCGGTTTACCAAGCGCTCCATATCATGTGAAAGCC 1080
 DB 8888 CAGCAAGCAAGCTGTTCCATCGGTTTACCAAGCGCTCCATATCATGTGAAAGCC 8947
 QY 1081 GATTTTCATGAGCGAGATATTTGTTGGCTTCCGCTGCTTCAATGAGCGCTTCCCAAT 1140
 DB 8948 GATTTTCATGAGCGAGATATTTGTTGGCTTCCGCTGCTTCAATGAGCGCTTCCCAAT 9007
 QY 1141 TCAGACACATTAAGTGGGCTCATGACTCTGCGCAATTTGGCTGAAACCGTTTGGGAG 1200
 DB 9008 TCAGACACATTAAGTGGGCTCATGACTCTGCGCAATTTGGCTGAAACCGTTTGGGAG 9067
 QY 1201 GATTCCTCTCTCTCAGGAGTAAAGAGGATCTTCTGGGCAACATCTCTGCTGAG-TAA 1259
 DB 9068 GATTCCTCTCTCTCAGGAGTAAAGAGGATCTTCTGGGCAACATCTCTGCTGAGGAA 9127
 QY 1260 ACATGAGTTTCTCTCATGATGAGGACCTTATTCAGATTAATGGAATAGATCTGGCGG 1319
 DB 9128 ACATGAGTTTCTCTCATGATGAGGACCTTATTCAGATTAATGGAATAGATCTGGCGG 9186
 QY 1320 GTTTGGTCAAGTTTGAAGCGGTTTATTAAGATCTCTCGCTGCTCAACGAGTA 1379
 DB 9187 GTTTGGTCAAGTTTGAAGCGGTTTATTAAGATCTCTCGCTGCTCAACGAGTA 9246
 QY 1380 TGAAGAAATACAGCGGATGTGCTGAGAGAAATCTCGAATCTTCAAGTGGATCTGCATC 1439
 DB 9247 TGAAGAAATACAGCGGATGTGCTGAGAGAAATCTCGAATCTTCAAGTGGATCTGCATC 9306
 QY 1440 TGAAGTGTAAAGGTGTCTGAGAGCCAGCGATATGCGATGTTCAAGTGGAGCGAT 1499
 DB 9307 TGAAGTGTAAAGGTGTCTGAGAGCCAGCGATATGCGATGTTCAAGTGGAGCGAT 9366
 QY 1500 TCAGAAAGAAAGCAAAATTAAGATTAAGGCTTAAGAGCGGATATCTGAACCTTAATGA 1559
 DB 9367 CGAGAAAGAAAGCAAAATTAAGATTAAGGCTTAAGAGCGGATATCTGAACCTTAATGA 9426
 QY 1560 TAAGGTGTGTCAACTGTGAGCTCGCAATATCTCAAGTGGATCTGCATGCAATGCA 1619
 DB 9427 TAAGGTGTGTCAACTGTGAGCTCGCAATATCTCAAGTGGATCTGCATGCAATGCA 9486
 QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACGCAATGACAGATC 1679
 DB 9487 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACGCAATGACAGATC 9546
 QY 1680 GTCAAAATCTTCTCTGATGACTGAACGAAATTTCTGTTAGACCAATCTCTCCGCTTG 1739
 DB 9547 TTTCAAAATCTTCTCTGATGACTGAACGAAATTTTCTGTTAGACCAATCTCTCCGCTTG 9606
 QY 1740 TGTCTCTATGAGCGGATTCGCAAAAGCAAGTGTATGCTCTGAGTATAGTTCGAGATCC 1799
 DB 9607 TGTCTCTATGAGCGGATTCGCAAAAGCAAGTGTATGCTCTGAGTATAGTTCGAGATCC 9666

QY 1800 GA 1801
 DB 9667 GA 9668
 RESULT 15
 AVU83986
 LOCUS
 DEFINITION
 AVU83986 14960 bp DNA linear BCT 02-MAR-2001
 Agrobacterium vitis plasmid pTIG474 agrocinopine synthase, (5'),
 and (1aah) genes, (1aah) pseudogene, (lpc) and (6b) genes, and
 (lpc) pseudogene, complete cds.
 ACCESSION
 U83986
 U83986.1 GI:1814315
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Agrobacterium vitis (Rhizobium vitis)
 Agrobacterium vitis
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 REFERENCE
 1 (bases 1 to 14960)
 Otten, L. and De Ruifray, P.
 Sequence and functional analysis of the shoot-inducing octopline
 T-DNA of Agrobacterium vitis strain CG474
 Unpublished
 2 (bases 1 to 14960)
 Otten, L. and De Ruifray, P.
 Direct Submission
 Submitted (07-JAN-1997) Institute for Plant Molecular Biology, rue
 du General Zimmer 12, Strasbourg 67084, France
 FEATURES
 source
 1..14960
 /organism="Agrobacterium vitis"
 /mol_type="genomic DNA"
 /strain="CG474"
 /db_xref="taxon:373"
 /plasmid="pTIG474"
 7244..8395
 /organism="Agrobacterium vitis"
 /mol_type="genomic DNA"
 /db_xref="taxon:373"
 /insertion_seq="15870-like insertion element"
 11921..12767
 /organism="Agrobacterium vitis"
 /mol_type="genomic DNA"
 /db_xref="taxon:373"
 /insertion_seq="15869"
 252..14745
 /note="shoat-inducing octopline T-DNA"
 252..278
 /note="flanking repeat; left border of the T-DNA"
 /rpt_type=direct
 complement(496..1854)
 /codon_start=1
 /transl_table=11
 /product="agrocinopine synthase"
 /protein_id="AAB41866.1"
 /db_xref="GI:1814316"
 /translation="MALNRSVLYLLTMTCHNIPFFPAQVLCKRKHVETGVPAMT
 LHMDELBNVTPLDVLARHSVPLRLLEEDLPLQVTEHGMELGRIOECTSLLA
 GAGQBSRNLSMDVCTSDNVALVSDLTMTWVSELDGKLPETHSSTIKQVTVIR
 EVSNGTIQKRYFEVDHIPSLEILSKVFLANPDATFDGRNYEAAHVYAMSHRPE
 YHORVVVLLVTFEYPRGAFVDAVLAQPSWRSIALMPALFPEELCRLARLVOT
 EPTVDLDYLAGRAMIDSMLODMRIVAHHVSVGTVRNLDGVDDVLAAPDSDEAA
 VRLAYVKEDTMRAPKRPKFAAVRCYDFALLDCCGRGSEFSIDIKTGRRARHETD
 ERKHTRMRKGTGNSATIDWVLSDRPEDEMAIWEWRNGIDREVSHTLSPLDVSVET
 PG"
 2363..3046
 /gene="5"
 /gene="5"
 2363..3046
 /codon_start=1
 /transl_table=11
 /protein_id="AAB41867.1"
 /db_xref="GI:1814317"

```
/translation="MYHSRPIFNVIDSSNIGHREDKFLRHVEIAYRSPARDLIRA
ORSLINIIIDVIDPAIDEVNRFCEVACLOGPAGIPLTIIILNDSIMVYCSFEM
RRAHKLRYDVGDSVIVSTVPYTEGIRKETMRSSNONNVGENTDEIHNVAYIAP
LPTSLONPSVSHMKIRCNSEFLAPSRVDPCEVBIIVAGKALFEDNGLKKPKMGEMA
LTSIMKRLV"
gene
complement(3569..4972)
/gene="iaah"
/gene="iaah"
complement(3569..4972)
/codon_start=1
/transl_table=1
/protein_id="AAB41868.1"
/db_xref="GI:1814318"
/translation="MYSITSLAOSLEHLRRKDYSCLEVETLIARCEAAKSLNALLAT
DMGLRSAAKIDRHNGAGVGLGIPCEFKANIATGVPPSATPAILNHLPIKIPSRV
AERLPSAGALPGASGNMHELSFGITSNNTATGAVRPNPNSLISGSSGVAAYASR
LMJGIGITGASVRLPALCGVGFRTLGRIPRRIIPVSTRTAGIACVADV
ILIDQVLSGRPARILPILKGLRIGLPTTFYDDLDADAVAFATRIILANNGVTFV
EADIPLHEDXXGASLPIALYEFPAKQYLDPLVDGADVINGIRSDPANIYKA
QVNDHQISSDEYELAROSFRLQATYRNFRLYQDALIFPTAPLAARAKIGDFSVT
HNGSTMTFELIYVRNDPSSNAGLPGSLPVLHTPRRLPGMEIDDLASDRHLAIG
AALEKADITHLPFNVLN"
gene
5306..8720
/gene="iaam"
join(5306..7243,8396..8720)
/gene="iaam"
/translation="interrupted by a bacterial IS870-like insertion
element"
/pseudo
/codon_start=1
/transl_table=1
/repeat_region
7244..7247
/translation="flanking repeat; left border of the IS870-like
insertion element"
/tpc_type=direct
/gene="ORF1"
7301..8395
/gene="ORF1"
/gene="ORF1"
/codon_start=1
/transl_table=1
/protein_id="AAB41869.1"
/db_xref="GI:1814319"
/translation="MRIMCWRGTGVSFTVSPDTHORIMALTDRDNAPOKHVPSEILL
SADGCVTRIMROTKSKTCVMRQERPAEGEGLLRKTRESRIYRLGPDYAEVYV
ARRILIRARROQMTADMMATVAGISARIRIWKAGLAIPHRWQPKLSNPKFYDK
LRDVGVLVDPRAHALVLSVDEKSOIQALDRTPGLPMKGRIGTHTHYKSGITTL
PALANVLDGTVIGCNNQRHROHFIPLKRIKINAQADKSVHVLIDNVAHAKPKVRA
WLEHREFTFHFPTSCSWLNAVEGFARLSKRRLKRGVFSHVVDLQAAINRFLVEHN
KKRPFMTADPPKIIAAVKGHQAALDSIH"
repeat_region
8392..8395
/translation="flanking repeat; right border of the IS870-like
insertion element"
/tpc_type=direct
/gene="ipc"
9419..10141
/gene="ipc"
/gene="ipc"
/codon_start=1
/transl_table=1
/protein_id="AAB41870.1"
/db_xref="GI:1814320"
/translation="MDRLIRFGPTCTGKSTAIRLAQOQGLPVLSIDRVCCPOLSTG
SGKPTVEELKGTIRLYLEDRPLVKGIIIAKQAHRLIGEVNTEAGSLILBGSISL
LRCMAQSSYSTWDFRWHLIRKLADETEPMNAKARVQMLRPAVGPSTIOELVHLNM
ERLRLILKEIDGRYRAMLFAEQNOITPMLQLDPMDEGELIHGIAQSVYLLIHAROE
OEFPVSVVAFEFEGPFGMC"
gene
complement(10816..11439)
/gene="6b"
complement(10816..11439)
/codon_start=1
/gene="6b"
/transl_table=1
/codon_start=1
/transl_table=1
```

```
/protein_id="AAB41871.1"
/db_xref="GI:1814321"
/translation="MTVANVOYRDLITLISAGEIOSRLRQARTDREGALLGDIYVQPS
VLSGHPDEBYIFSQELVYVYLREDIAQCRLRURPENSNSNFGVATRAITPRMLNA
HGLNVLQRCDPGGVNVYGGPHATNPOFLAIRSCVFRFASVDVNNSSYGYARNG
NYTERAGGGDDVDEGEAATAATPESQIGNLINPIVALGSCQLSA"
gene
complement(13214..14290)
/gene="oc8"
complement(13214..14290)
/gene="oc8"
/codon_start=1
/transl_table=1
/repeat_region
14719..14745
/translation="flanking repeat; right border of the T-DNA"
/tpc_type=direct
BASE COUNT 3854 a 3510 c 3632 g 3958 t 6 others
ORIGIN
Query Match 81.3%; Score 1464.4; DB 1; Length 14960;
Best Local Similarity 89.0%; Pred.No.0;
Matches 1604; Conservative 0; Mismatches 196; Indels 2; Gaps 2;
1 ATGTCAGCTTCACTCTCTCTGATTAACGATGATCATCTCCCAACAAATGATGAT 60
5306 ATGTACCTTCACTCTCTCTGATTAACGATGATCATCTCTACCAAAATGATGAT 5365
61 CTGACATATGATGATTAAGCGGATGATTAAGCAACGAGGATTCGATCTTCTTGA 120
5366 CTGATATATGATGATTAAGCGGATGATTAAGCAACGAGGATTCGATCTTCTG 5425
121 CGAGAACTTCAAGGGAAGAGAGATTACTCAAGCTCCACGAGTGCAGCGCTGGGTA 180
5426 CGAGAACTTCAAGGGAAGAGAGATTACTCAAGCTCCACGAGTGCAGCGCTGGGTA 5485
181 GCTTGCAAAAGGCTGCGCATGATGATGATGATGATGATGATGATGATGATGAT 240
5486 GCTTGCAAAAGGCTGCGCATGATGATGATGATGATGATGATGATGATGATGAT 5545
241 GCTTCTCTCGCTTATATCTATATGGAAGAAATTTGGGCGGATTAATGATGA 300
5546 GTCTCTCTCGCTTATATCTATATGGAAGAAATTTGGGCGGATTAATGATGA 5605
301 CCTTGCGCGCGGCAACAGTAGTGTCTGTCATGATGATGATGATGATGATGATG 360
5606 CCATGCGCGCGGCAACAGTAGTGTCTGTCATGATGATGATGATGATGATGATG 5665
361 GATTTCTCGAAGCAACATACTCAAGCGCTGTTTCTGAGCGGTAAAGATGCA 420
5666 GATTTCTCGAAGCAACATACTCAAGCGCTGTTTCTGAGCGGTAAAGATGCA 5725
421 CGGATGATCTTATGATCTTATGATGATGATGATGATGATGATGATGATGATG 480
5726 CCAGCGACCTTATGATCTTATGATGATGATGATGATGATGATGATGATGATG 5785
481 CTGCAATGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
5786 CTGCAATGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5845
541 CTGCAAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
5846 CTGCAAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5905
601 TCGGAGGCTCTTCAACATGATGATGATGATGATGATGATGATGATGATGATG 660
5906 TCGGAGGCTCTTCAACATGATGATGATGATGATGATGATGATGATGATGATG 5965
661 TGTTCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
5966 TGTTCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6025
721 GTCAATGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
```

Search completed: November 24, 2003, 01:37:14
Job time : 6650.53 secs

```
Db 6026 GTCATTGGGGCTGGCAATTCGGGCTCTGTCGTGTCGAAGCGACTGCTTCATGCTGTGTA 6085
QY 781 GAGGATGTTACATATATATGAAGCAAGTATCGTGTGGAGGCAAGCTTGTGTCATGCT 840
Db 6086 GACGATGTTACATATATGAAGCAAGTATCGGTTGAGGCAAGCTTGTGTCATGCT 6145
QY 841 TTCAGGAGCGCTCTGATGTCGTGCGGAATGGGGCGATGCGATTTCTCTGCTGCA 900
Db 6146 TTCAGGAGACTCGAGCGTCGAAGCGGAATGGAGCGATGGCATTTCTCTGCTGCA 6205
QY 901 TTCGCTGTTGTTTCTCTCTGAGCGTTACGCGCTGTCTTGATGAGGCCGTTCCCAAT 960
Db 6206 TCGTGTCTGTTCTTCTCTCTCGAGCGGTACGCTGTCTTGATGAGGCCATTCCTCAAT 6265
QY 961 CCGGCAACAGTACACTTCTGCTGTCACCAAGCGCTCCAATACATGTGGAAGCGGG 1020
Db 6266 CCGGCAACAGTACACTTCTGCTGTCACCAAGCGGTCCAATACATGTGGAAGCGGG 6325
QY 1021 CAGCTGCACCGAAGCTGTTCCATCCGCTTAAACAACGTTGGCGTCCGTTCTGAAAGAC 1080
Db 6326 CAGCCGCCACCGAAGTGTTCATCCGCTTAAACAACGGTGGCATGCTTTCTTAAAGAC 6385
QY 1081 GGTTCATGAGCGAGATATGTTGCTTCCGCTGTCCCTATTAATCAAGCCTTGAA 1140
Db 6386 GGTTCCTTGAAGGAGATATGTTGCTTCCGCTGTATCACTGAGGCCCTGAAA 6445
QY 1141 TCAGGACACATTAGTGGGCTCATGACTCTGGCAATTTGGCTGAACGTTCCGGAGG 1200
Db 6446 TTAGGGACATTAAGCGGGCTCATGACTCTGGCAATTTGGCTGAACGTTCCGGAGG 6505
QY 1201 GAGTCTCTCTTCAGGATAGAGAGATCTTCTGGGCAACATCTCTGTGTGT-AA 1259
Db 6506 GAGTCTCTCTTCAGGATAGAGAGATCTTCTGGGCTGCAATCTCCGGTGTGAA 6565
QY 1260 ACATGAGTTTCTCTCATGATGGGACCTATTCAGCTAAATGGGAATAGATCTGGCGG 1319
Db 6566 ACATGAGTTTCTCTCATGATGGGACCTATTCAGCTAAATGGGAATAGATCTGGC-GG 6624
QY 1320 GTTTGTCAGTTTAAAGCGGGTTATTGAGATCTCGCGCTGGTCAATCAACGGATA 1379
Db 6625 GTTTGTCAGTTTAAAGCGGGTTATTGAGATCTCTGCTGTGATTAACGGATA 6684
QY 1380 TGAAGAAATCAGCGGATGTGCCCTGAAGGATCTCAGAACTTCCAGCTCGATCGCATC 1439
Db 6685 TGAAGAAATCAGCGGATGTGCTCTGAAGGATCTCAGAACTTCCAGCTCGATCGCATC 6744
QY 1440 TGAAGTGTAAACGCTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAAGCGCAT 1499
Db 6745 TCAAGTGTAAATGTGTGTCTGTAAACGCGTATACGCCATGTTCAGATTAGGCGCAT 6804
QY 1500 TCAGAAAGAAAGCAAAATTAAGATTAAGGCTTAAGAGCGGGATCTGAACCTTATGA 1559
Db 6805 TGAGAAAGAAAGCAAAATTAAGATTAAGGCTTAAGAGCGGGATCTGAACCTTATGA 6864
QY 1560 TAAAGTGTGTCACTGTGACTCGCAAAATATCAACTCAGGCAATTCCTGACATGCGA 1619
Db 6865 TAAAGTGTGTCACTGTGACTCGCAAAATATCAACTCAGGCAATTCCTGACATTCGA 6924
QY 1620 TACCAATATTTTCAGGCAACAGTGAACAAGCGGTTGATPAACGCCATATGACAGATC 1679
Db 6925 TACCAATATTTTCAGGCAACAGTGAACAAGCGGTTGATPAACGCCATATGACAGATC 6984
QY 1680 GTCAAAACTCTTCTGTATGACTGAACGAAATTTCTGTTAGACCATATCTCTCCGTCTG 1739
Db 6985 TTCAAATATTTTGTGCTGACTGAACGAAATTTTGTGATCAACGATATGCTCCGCTCTG 7044
QY 1740 TGTCTCATGACGAGATTCGCAAAAGAGTATTTGCTGGAATGATGTCAGAGATCC 1799
Db 7045 TGTCTCATGACGAGATTCGCAAAAGAGTATTTGCTGGAATGATGACCGAAGATCC 7104
QY 1800 GA 1801
Db 7105 GA 7106
```


Db 1801 A 1801

RESULT 2

AAD00625 standard; DNA; 1807 BP.

AAD00625;

29-AUG-2000 (first entry)

Agrobacterium tumefaciens T-DNA mutant iaam gene.

iaam; iaah; ipt; T-DNA; transferred DNA; gall-disease resistance; BR; bacterial resistance construct; untranslatable RNA; tumour; oncogene; auxin; plant growth hormone; mutant; ds.

Agrobacterium tumefaciens.

W0200026346-A1.

11-MAY-2000.

04-NOV-1999; 99WO-US26100.

05-NOV-1998; 98US-0107185.

(UYOR-) UNIV OREGON STATE.

Ream W, Mok MC, Lee H;

WPI; 2000-365599/31.

Gall resistant plants generated by transforming cells with an untranslatable nucleic acid homologous to a gall disease-causing gene

Example C; Page 43-44; 4ppp; English.

The patent discloses a method for producing plants resistant to gall disease, induced by Agrobacterium infection, by transformation with a bacterial resistance (BR) construct capable of eliciting co-suppression. The construct encodes an untranslatable RNA molecule, that is highly homologous to the tumour or gall disease-causing genes iaam, iaah and ipt of Agrobacterium tumefaciens. These oncogenes are modified by introduction of premature termination codons or frameshift mutations, to inhibit their expression. This method is used to reduce susceptibility of plants to gall diseases. The present DNA sequence is the Agrobacterium tumefaciens T-DNA (transferred DNA) mutant iaam gene. This oncogene was modified by introduction of a stop sequence (TGA) at the third codon and deletion of two bases following the third codon. This sequence was used for the construction of double stranded RNA and untranslatable dsRNA vectors, that reduce gall formation in transformed plant cells.

Sequence 1807 BP; 428 A; 418 C; 486 G; 475 T; 0 other;

Query Match 99.3%; Score 1788; DB 21; Length 1807;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

12 ACCCTCTCTGATACACGATCGGATCATCTCCCAACCAAAATGGTGAATCGACAAATGT 71
18 ACCCTCTCTGATACACGATCGGATCATCTCCCAACCAAAATGGTGAATCGACAAATGT 77
72 CGATTAAGCGGATGAATTTGACCGCAGGGTTTCGATGCCCTTTAGAACGAGAAGCTTC 131
78 CGATTAAGCGGATGAATTTGACCGCAGGGTTTCGATGCCCTTTAGAACGAGAAGCTTC 137
132 TAGGGGAGAGAGATTAATCAAGCTTCACCGAGTGCAGCGCTGGTTAGCTTGCAAAAG 191
138 TAGGGGAGAGAGATTAATCAAGCTTCACCGAGTGCAGCGCTGGTTAGCTTGCAAAAG 197

QY 192 GCTGCGGATGTCGCTTCCCGAGATCTCAGCTGTGGAAAGGTACAGTTCTCCGC 251
DB 198 GCTGCGGATGTCGCTTCCCGAGATCTCAGCTGTGGAAAGGTACAGTTCTCCGC 257
QY 252 TTATATCTAATATGCGCAAGAAATTTCTGGGCGGATACCTTGAATCGAACTTTGGGCGG 311
DB 258 TTATATCTAATATGCGCAAGAAATTTCTGGGCGGATACCTTGAATCGAACTTTGGGCGG 317
QY 312 GCGAAGAGTGAAGTGTCTGCTTCCGATCGAATCGAATCGAATCGAATCGAATCGAATCGA 371
DB 318 GCGAAGAGTGAAGTGTCTGCTTCCGATCGAATCGAATCGAATCGAATCGAATCGAATCGA 377
QY 372 AGCACAATTAATCCAGCCCTGTTTGTCTGAGCGGTAAGAAAGATGTGCAACGATTAATCT 431
DB 378 AGCACAATTAATCCAGCCCTGTTTGTCTGAGCGGTAAGAAAGATGTGCAACGATTAATCT 437
QY 432 TAGCATTTTGTGGCCATTTCAATCTCTAAGACTGCGGCTTTGAAACCTTGCCCAATGCC 491
DB 438 TAGCATTTTGTGGCCATTTCAATCTCTAAGACTGCGGCTTTGAAACCTTGCCCAATGCC 497
QY 492 GCTGTAGAGAAATGCGACGATGAAATGCCCTTACCGGGTTACCATACCTTTGAAGGGGC 551
DB 498 GCTGTAGAGAAATGCGACGATGAAATGCCCTTACCGGGTTACCATACCTTTGAAGGGGC 557
QY 552 CGTGCCATTTGACATGATGATATGCTGCGAAACCTGATGCTGAAGGGTTCCGAGGTTCC 611
DB 558 CGTGCCATTTGACATGATGATATGCTGCGAAACCTGATGCTGAAGGGTTCCGAGGTTCC 617
QY 612 CTTTCCAAATCGACTTGTCTACGACTGCGAAGCGTTTGTGACCAATGTTCCGATAG 671
DB 618 CTTTCCAAATCGACTTGTCTACGACTGCGAAGCGTTTGTGACCAATGTTCCGATAG 677
QY 672 TGACGAGATGCGCTTCTTCCGAGGATGTTCTTAAGCCGAAGTGGCGGTCAATGGCGC 731
DB 678 TGACGAGATGCGCTTCTTCCGAGGATGTTCTTAAGCCGAAGTGGCGGTCAATGGCGC 737
QY 732 TGGCATTTCCGACCTGCTGTGGCAACGAACGCTTCATGCTGGGTAGACGATGTAC 791
DB 738 TGGCATTTCCGACCTGCTGTGGCAACGAACGCTTCATGCTGGGTAGACGATGTAC 797
QY 792 AATATATGAGCAATGATGCTGTGGAGCAAGCTTTGCTCAATGCTTTCAAGGAGCG 851
DB 798 AATATATGAGCAATGATGCTGTGGAGCAAGCTTTGCTCAATGCTTTCAAGGAGCG 857
QY 852 TCTAGTGTGTGCGCGAAATGGGGGCGATGCGATTTCTCTGCTGCTGCTGCTGCTGCT 911
DB 858 TCTAGTGTGTGCGCGAAATGGGGGCGATGCGATTTCTCTGCTGCTGCTGCTGCTGCT 917
QY 912 TTTCTTCTGAGCGTTAAGCGCTGCTTCTGATGAGCGGCTTCCAAATGCCGCGACAGT 971
DB 918 TTTCTTCTGAGCGTTAAGCGCTGCTTCTGATGAGCGGCTTCCAAATGCCGCGACAGT 977
QY 972 CGACACTTAATGCTGTACCAAGCGCTCAATATCATGTGGAAGCCGGGAGCTGCCACC 1031
DB 978 CGACACTTAATGCTGTACCAAGCGCTCAATATCATGTGGAAGCCGGGAGCTGCCACC 1037
QY 1032 GAAGCTGTTCAATGCGGTTTCAACGCGTGGCGCTGCTTTGAAGAGCGGTTTCAATGA 1091
DB 1038 GAAGCTGTTCAATGCGGTTTCAACGCGTGGCGCTGCTTTGAAGAGCGGTTTCAATGA 1097
QY 1092 GCGAGATATGATGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1151
DB 1098 GCGAGATATGATGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1157
QY 1152 TAGTGGGCTCATGATCTCTGCGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1211
DB 1158 TAGTGGGCTCATGATCTCTGCGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1217
QY 1212 TTGAGGATATGAGGATCTTTCTGGGCAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1271
DB 1218 TTGAGGATATGAGGATCTTTCTGGGCAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1277
QY 1272 CTTCAATGATTTGGGACCTATTCAAGCTAATGGGAATGGAATCTGGCGGGGTTGGTCAAGT 1331

```
Db 1278 CCTCATATTGGGACCTATTCACGCTAAATGGGAATAGGATCTGGCGGGTTTGCTCAAGT 1337
Qy 1332 TTTTGAAGCGGGTTTATGAGATCTCCGCTTGTCATCAAGGATATGAAGAAATCA 1391
Db 1338 TTTTGAAGCGGGTTTATGAGATCTCCGCTTGTCATCAAGGATATGAAGAAATCA 1397
Qy 1392 GCGGATGTGCGCCGTAAGGAATCTCAGAACTTCGACGTCGATTCGCAATGAGTGTAA 1451
Db 1398 GCGGATGTGCGCCGTAAGGAATCTCAGAACTTCGACGTCGATTCGCAATGAGTGTAA 1457
Qy 1452 CGGCTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGCGCATTCAGAGAAAA 1511
Db 1458 CGGCTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGCGCATTCAGAGAAAA 1517
Qy 1512 GACAAAATAAAGATAAGCTTAAAGACCGGAGATCTGAACTTTATGATAAAGTGCTGT 1571
Db 1518 GACAAAATAAAGATAAGCTTAAAGACCGGAGATCTGAACTTTATGATAAAGTGCTGT 1577
Qy 1572 CACATCTGGAAGTCCGAATATCCAACTCAGGCAATGGCTGACATGCGATACCAATATTTT 1631
Db 1578 CACATCTGGAAGTCCGAATATCCAACTCAGGCAATGGCTGACATGCGATACCAATATTTT 1637
Qy 1632 TCAGGCAACAGTGAACCAAGCGGTTGATTAACAGGCGCATATGACAGATCGTCAAACTCTT 1691
Db 1638 TCAGGCAACAGTGAACCAAGCGGTTGATTAACAGGCGCATATGACAGATCGTCAAACTCTT 1697
Qy 1692 CCGTATGACTGAACGAAATTTCTGTTAGACCATATCTCCGCTTGCTGTCTCATGA 1751
Db 1698 CCGTATGACTGAACGAAATTTCTGTTAGACCATATCTCCGCTTGCTGTCTCATGA 1757
Qy 1752 CGGATGCGCAAAAGCACTGATTCGCTGAGTATGAGTATGCGAGATCC 1799
Db 1758 CGGATGCGCAAAAGCACTGATTCGCTGAGTATGAGTATGCGAGATCC 1805

RESULT 3
AAD00626
ID AAD00626 standard; DNA; 2544 BP.
AC AAD00626;
XX
DT 29-AUG-2000 (first entry)
XX
DE Agrobacterium tumefaciens ipt/iaam untranslatable BR construct.
XX
KM iaam; iaah; ipt; T-DNA; transferred DNA; gall-disease resistance; BR;
KM bacterial resistance construct; untranslatable RNA; tumour; oncogene;
KM plant growth hormone; ds.
XX
OS Agrobacterium tumefaciens.
XX
FH Key 1..309 Location/Qualifiers
FT misc_feature /*tag= a
FT /*note= "Corresponds to Agrobacterium mutant ipt gene"
FT /*tag= b
FT /*note= "Corresponds to Agrobacterium mutant iaam gene"
XX
PN WO200026346-A1.
XX
PD 11-MAY-2000.
XX
PF 04-NOV-1999; 99WO-US26100.
XX
PR 05-NOV-1998; 98US-0107185.
XX
PA (UYOR-) UNIV OREGON STATE.
XX
PI Ream W, Mok MC, Lee H;
XX
DR WPI; 2000-365599/31.
```

```
XX Gall resistant plants generated by transforming cells with an
PT untranslatable nucleic acid homologous to a gall disease-causing gene
PT
XX
PS Example c; Page 44-45; 49pp; English.
XX
CC The patent discloses a method for producing plants resistant to gall
CC disease, induced by Agrobacterium infection, by transformation with a
CC bacterial resistance (BR) construct capable of eliciting co-suppression.
CC The construct encodes an untranslatable RNA molecule, that is highly
CC homologous to the tumour or gall disease-causing genes iaam, iaah and
CC ipt of Agrobacterium tumefaciens. These oncogenes are modified by
CC introduction of premature termination codons or frameshift mutations, to
CC inhibit their expression. This method is used to reduce susceptibility
CC of plants to gall diseases. The present DNA sequence is the Agrobacterium
CC tumefaciens bacterial resistance construct, that encodes untranslatable
CC RNA forms of both the mutant ipt and iaam genes. This construct can
CC suppress both shooty and necrotic gall formation in transgenic plants.
XX
SQ Sequence 2544 BP; 625 A; 592 C; 674 G; 653 T; 0 other;
XX
Query Match 99.3%; Score 1788; DB 21; Length 2544;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12 ACCTTCCTTGATTAACCAAGTGCATATCTCCCAACCAAAATGATGATCTGCAATGCT 71
Db 318 ACCTTCCTTGATTAACCAAGTGCATATCTCCCAACCAAAATGATGATGCAATGCT 377
Qy 72 CGATTAAGGCGGATGAATTTGACCGCAGGTTTCCGATGCTCTTGAACGAGAAAGCTTC 131
Db 378 CGATTAAGGCGGATGAATTTGACCGCAGGTTTCCGATGCTCTTGAACGAGAAAGCTTC 437
Qy 132 TAGGGGAAGGAGATTACTCAAAAGCTCCACGAGTCAAGCGCTGGTTAGCTTGGCAAAAG 191
Db 438 TAGGGGAAGGAGATTACTCAAAAGCTCCACGAGTCAAGCGCTGGTTAGCTTGGCAAAAG 497
Qy 192 GCTGGCGGATGTCGCTTCCCGAGATCTCAGTGTGTGAAAGGTAGCATTTCTCCGCG 251
Db 498 GCTGGCGGATGTCGCTTCCCGAGATCTCAGTGTGTGAAAGGTAGCATTTCTCCGCG 557
Qy 252 TTATATCTATATTTGGGAAAGAAATTTCTGGGGCGGATCTTGAATTCGAACCTTTGGGCGG 311
Db 558 TTATATCTATATTTGGGAAAGAAATTTCTGGGGCGGATCTTGAATTCGAACCTTTGGGCGG 617
Qy 312 GGCACAGTGAAGTGTCTCGTTCGATCGACTTGAACACCATTTGATGATTTCTCCGA 371
Db 618 GGCACAGTGAAGTGTCTCGTTCGATCGACTTGAACACCATTTGATGATTTCTCCGA 677
Qy 372 AGCACAATATATCAAGCCCTGTTTTTGTGACGGGTAAAGATGTGCACCCGATTTGATCT 431
Db 678 AGCACAATATATCAAGCCCTGTTTTTGTGACGGGTAAAGATGTGCACCCGATTTGATCT 737
Qy 432 TAGTCATTTGCGGCCATTTCAATCTCTAAGATCGCGGCTTTGCAACCTGCAATGCGC 491
Db 738 TAGTCATTTGCGGCCATTTCAATCTCTAAGATCGCGGCTTTGCAACCTGCAATGCGC 797
Qy 492 GCTGTACGAAATGAGCAGATGAATGCGTTACCGGGTTTACCATTAACCTTTGAAGGGGC 551
Db 798 GCTGTACGAAATGAGCAGATGAATGCGTTACCGGGTTTACCATTAACCTTTGAAGGGGC 857
Qy 552 CGTGCCATTTGACATGTAGCTTATGCTGAACCTGTAGCTGAAGGTTCCGCAAGTTTC 611
Db 858 CGTGCCATTTGACATGTAGCTTATGCTGAACCTGTAGCTGAAGGTTCCGCAAGTTTC 917
Qy 612 CTTTCAACATGAGCTGTACAGATGCAAGCGTTTGTGACCAATGTTCCGATAG 671
Db 918 CTTTCAACATGAGCTGTGTACAGATGCAAGCGTTTGTGACCAATGTTCCGATAG 977
Qy 672 TGAGCGATCGCTTCTTTCGAGAGATGTTCTTAAGCCGAAAGTGCGGCTCATTTGCGCG 731
Db 978 TGAGCGATCGCTTCTTTCGAGAGATGTTCTTAAGCCGAAAGTGCGGCTCATTTGCGCG 1037
```



```

QY 732 TGGCATTTCCGACTCGTGGTGCAAAAGAACTGCTTCATGCTGGGTAGACGATGTTAC 791
DB 1038 TGGCATTTCCGACTCGTGGTGCAAAAGAACTGCTTCATGCTGGGTAGACGATGTTAC 1097
QY 792 AATATATGAAGCAAGTGAATGCTGTGGAGCAAGCTTTGGTCAATGCTTTTCAGGACGC 851
DB 1098 AATATATGAAGCAAGTGAATGCTGTGGAGCAAGCTTTGGTCAATGCTTTTCAGGACGC 1157
QY 852 TCGTATGCTGCTGGCGCAATGGGGGGGATGCGATTTCTCTGCTGCTGCAATTCGCTTTGTT 911
DB 1158 TCGTATGCTGCTGGCGCAATGGGGGGGATGCGATTTCTCTGCTGCTGCAATTCGCTTTGTT 1217
QY 912 TTTCTTCTCGAGCGTTACGGGCTGTCTTGCATGAGGCGGCTTCCCAAAATCCCGGACAGT 971
DB 1218 TTTCTTCTCGAGCGTTACGGGCTGTCTTGCATGAGGCGGCTTCCCAAAATCCCGGACAGT 1277
QY 972 CGACACTTACTTGGTGTACCAAGGCGTCCAATAATGTTGAAAAGCCGGGAGCTGCACCC 1031
DB 1278 CGACACTTACTTGGTGTACCAAGGCGTCCAATAATGTTGAAAAGCCGGGAGCTGCACCC 1337
QY 1032 GAAGCTGTCCATCCGCTTTTACACAGGTTGGCGTGGCTTTTGAAGGACGGTTTTCATGA 1091
DB 1338 GAAGCTGTCCATCCGCTTTTACACAGGTTGGCGTGGCTTTTGAAGGACGGTTTTCATGA 1397
QY 1092 GCGAGATATGTTGTTGGCTTCCGCTGCTGCTATTAATCAGGCGCTTGAATCAGACACAT 1151
DB 1398 GCGAGATATGTTGTTGGCTTCCGCTGCTGCTATTAATCAGGCGCTTGAATCAGACACAT 1457
QY 1152 TAGTGGGCTCATGACTCTCGGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1211
DB 1458 TAGTGGGCTCATGACTCTCGGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1517
QY 1212 TTACGGGATAGAGAGATCTTTCTGGGACACATCTCTCTGGTGTGAACATGAGATTTT 1271
DB 1518 TTACGGGATAGAGAGATCTTTCTGGGACACATCTCTCTGGTGTGAACATGAGATTTT 1577
QY 1272 CCTCATGATTGGGACCTATTCAGACTAATGGGAATAGGATCTGGCGGGGTTTGGTCAAGT 1331
DB 1578 CCTCATGATTGGGACCTATTCAGACTAATGGGAATAGGATCTGGCGGGGTTTGGTCAAGT 1637
QY 1332 TTTTGAAGCGGGTTTATGAGATCTCTCGCTTGTGATCAACGAGATATGAGAAATCA 1391
DB 1638 TTTTGAAGCGGGTTTATGAGATCTCTCGCTTGTGATCAACGAGATATGAGAAATCA 1697
QY 1392 GCGGATGTGCTCTGAAGGAATCTCAAGACTTCCAGCTCGGATGCAATCTGAAGTGTAA 1451
DB 1698 GCGGATGTGCTCTGAAGGAATCTCAAGACTTCCAGCTCGGATGCAATCTGAAGTGTAA 1757
QY 1452 CGGTGTGCTGTGAGCGCAGCGCATATGCCATGTTCAAGTCAAGGCGCATTCAGAAAGAAA 1511
DB 1758 CGGTGTGCTGTGAGCGCAGCGCATATGCCATGTTCAAGTCAAGGCGCATTCAGAAAGAAA 1817
QY 1512 GACAAAATTAAGATTAAGGCTTAAGACGGGATATCTGAACCTTATGATTAAGTGTGGT 1571
DB 1818 GACAAAATTAAGATTAAGGCTTAAGACGGGATATCTGAACCTTATGATTAAGTGTGGT 1877
QY 1572 CACATCTGGAATCTGCAAAATATCCAACTCAGGCAATGCTCTGACATGCGATACCAATATTTT 1631
DB 1878 CACATCTGGAATCTGCAAAATATCCAACTCAGGCAATGCTCTGACATGCGATACCAATATTTT 1937
QY 1632 TGAAGGACAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGGATTCGCAAAACTCTT 1691
DB 1938 TGAAGGACAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGGATTCGCAAAACTCTT 1997
QY 1692 CCTGATGACTGAACGAAAATTTCTGTTAGACATATCTCCGCTTGTGTTCTCTCATGA 1751
DB 1998 CCTGATGACTGAACGAAAATTTCTGTTAGACATATCTCCGCTTGTGTTCTCTCATGA 2057
QY 1752 CCGGATCGCAAAAGCAAGTGTATGCTGGAATATGAGTGGCAGAGATCC 1799
DB 2058 CCGGATCGCAAAAGCAAGTGTATGCTGGAATATGAGTGGCAGAGATCC 2105

```

```

RESULT 4
AAN50182
ID AAN50182 standard; DNA; 24596 BP.
XX
AC AAN50182;
XX
DT 25-MAR-2003 (updated)
DT 17-OCT-1991 (first entry)
XX
DE Complete nucleotide sequence of the T-DNA region of the octopine Ti
DE plasmid pTi15955.
XX
KW Plant vector; transformation-inducing principle (TIP) gene;
KW octopine Ti plasmid; ss.
XX
OS Agrobacterium tumefaciens ATCC 15955.
XX
PN EP14538-A.
XX
PD 19-JUN-1985.
XX
PF 16-NOV-1984; 84EP-0307969.
XX
PR 18-NOV-1983; 83US-0553786.
PA (AGRK ) AGRIGENETICS RES ASSOC LTD.
PA (LUBR ) LUBRIZOL GENETICS INC.
XX
PI Barker RF, Kemp JD;
PI
DR MPI; 1985-148223/25.
XX
XX New DNA vectors contg. T-DNA sequence of octopine Ti plasmid -
PT for expression in plant cells to confer desirable properties to
PT plants and their cells
XX
PS Claim 28; Fig 1; 87pp; English.
XX
CC The inventors claim a vector contg. a transformation-inducing
CC principle (TIP) gene from Ti plasmid pTi15955. The sequence of the
CC T-DNA of the octopine-type Ti plasmid has fourteen open reading
CC frames bounded by eukaryotic promoters, ribosome binding sites, and
CC polyadenylation sites. With the vectors, expression of structural
CC foreign genes in plant cells is promoted. The gene esp. encodes an
CC insecticidal toxin identical to or derived from the crystal protein
CC of Bacillus thuringiensis.
CC (updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 24596 BP; 6534 A; 5510 C; 5793 G; 6759 T; 0 other;

Query Match 98.0%; Score 1765.2; DB 6; Length 24596;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1792; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 1 ANGTCAGCTTCACTCTCTCTTATTAACGATGCGATCATCTCCCAACCAAAATGCTGAT 60
DB 5809 ATGTCAAGCTTCACTCTCTCTTATTAACGATGCGATCATCTCCCAACCAAAATGCTGAT 5868
QY 61 CTGACAAATGATGAAGCGGATGAATTGACCGGAGGTTTCCGATCCTTTCTTAGAA 120
DB 5869 CTGACAAATGATGAAGCGGATGAATTGACCGGAGGTTTCCGATCCTTTCTTAGAA 5928
QY 121 CGAGAACTTCTAGGGGAAAGAGATTACTCAAGCTCCACCGATGCGAGCGCTGGGTTA 180
DB 5929 CGAGAACTTCTAGGGGAAAGAGATTACTCAAGCTCCACCGATGCGAGCGCTGGGTTA 5988
QY 181 GCTTCCAAAAGCTGCGCATGCTGCTTCCCGGAGATCTCAGCTGCTGTGAAAAGTAGCA 240
DB 5989 GCTTCCAAAAGCTGCGCATGCTGCTTCCCGGAGATCTCAGCTGCTGTGAAAAGTAGCA 6048
QY 241 GTTCTCTCGCTTATATCTATATTGGAAGAAATTTCTGGGCGGATCTTGAATCGAAA 300

```

```

Db 6049 GTTCTCCGCTTATATCTATATTGGCAAGAAATTCGCGGCGGATTAATTGAAATGAAA 6108
QY 301 CCTTGGGCGGGGCAACAGTAGGTGCTGCTGCAATGCACTTGACAACCATTTGCATG 360
Db 6109 CTTTGGGCGGGGCAACAGTAGGTGCTGCTGCAATGCACTTGACAACCATTTGCATG 6168
QY 361 GATTTCTCCGAAACACAATACTAATCCAAGCCCTGTTTTGCTGAGCGGTAAAGATGCA 420
Db 6169 GATTTCTCCGAAACACAATACTAATCCAAGCCCTGTTTTGCTGAGCGGTAAAGATGCA 6228
QY 421 CCGATTGATCTTATGATCATTTGCTGGCCATTTCAATCTTAAGACTGCGGCTTTGAAAC 480
Db 6229 CCGATTGATCTTATGATCATTTGCTGGCCATTTCAATCTTAAAGACTGCGGCTTTGAAAC 6288
QY 481 CTGCAATGCGGCTGTAAGCAAGATGAGCAAGTAAGTAAGGCTTACCGGGTTACATTAAC 540
Db 6289 CTGCAATGCGGCTGTAAGCAAGATGAGCAAGTAAGTAAGGCTTACCGGGTTACATTAAC 6348
QY 541 CTTGAAGGGGCGGTGCAATTTGACATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 6349 CTTGAAGGGGCGGTGCAATTTGACATGATGATGATGATGATGATGATGATGATGATGAT 6408
QY 601 TCGGCAAGTTCTTTTCCAAATGCAATGCTCTAAGCACTGACAGCCGTTTTTGACCA 660
Db 6409 TCGGCAAGTTCTTTTCCAAATGCAATGCTCTAAGCACTGACAGCCGTTTTTGACCA 6468
QY 661 TGTTCGATGTGAGACGATGCGGCTCTTCCGAGGATGTTCTTAAAGCGGAAGTGGCG 720
Db 6469 TGTTCGATGTGAGACGATGCGGCTCTTCCGAGGATGTTCTTAAAGCGGAAGTGGCG 6528
QY 721 GTCAATGCGGCTGCAATTTCCGAGCTGATGATGATGATGATGATGATGATGATGATGAT 780
Db 6529 GTCAATGCGGCTGCAATTTCCGAGCTGATGATGATGATGATGATGATGATGATGATGAT 6588
QY 781 GACGATGTTCAATATATGAGCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 6589 GACGATGTTCAATATATGAGCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 6648
QY 841 TTGAGGAGCGCTCTGATGCTGCTGCGGCAATGGGGCGATGATGATGATGATGATGATGAT 900
Db 6649 TTGAGGAGCGCTCTGATGCTGCTGCGGCAATGGGGCGATGATGATGATGATGATGATGAT 6708
QY 901 TTCTGCTGTTTTTTCTCTCGAGCGTTACGCGCTGCTTCCGATGAGCCGTTCCCAAT 960
Db 6709 TTCTGCTGTTTTTTCTCTCGAGCGTTACGCGCTGCTTCCGATGAGCCGTTCCCAAT 6768
QY 961 CCGGCAAGTCACTTAATTGCTTCAAGCGGTCAATATGATGATGATGATGATGATGATGAT 1020
Db 6769 CCGGCAAGTCACTTAATTGCTTCAAGCGGTCAATATGATGATGATGATGATGATGATGAT 6828
QY 1021 CAGCTGCAACGAACTGCTTCAATCGGCTTCAACGCTTGGGCTGCTTGAAGGAC 1080
Db 6829 CAGCTGCAACGAACTGCTTCAATCGGCTTCAACGCTTGGGCTGCTTGAAGGAC 6888
QY 1081 GGTTCATGAGGAGATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 6889 GGTTCATGAGGAGATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6948
QY 1141 TCAGACACATTAAGTGGCTCATGATCTCTGCAAAATTTGGCTGAACGCTTTGGGAGG 1200
Db 6949 TCAGACACATTAAGTGGCTCATGATCTCTGCAAAATTTGGCTGAACGCTTTGGGAGG 7008
QY 1201 GAGTCCCTCTCTTCAAGGATAGAGAGATCTTCTGGGCAACATCTCTCGTGGGT-AA 1259
Db 7009 GAGTCCCTCTCTTCAAGGATAGAGAGATCTTCTGGGCAACATCTCTCGTGGGTGAA 7068
QY 1260 ACATGAGTTTCTCTCATGATGAGGACTTATCAAGCTAATGGAATAGATCTGCGGG 1319
Db 7069 ACATGAGTTTCTCTCATGATGAGGACTTATCAAGCTAATGGAATAGATCTGCGG-GG 7127
QY 1320 GTTTGATCAAGTTTGAACCGGGTTTATGAGATCTTCCGCTTGTGTCATCAACGGATA 1379
Db 7128 GTTTGATCAAGTTTGAACCGGGTTTATGAGATCTTCCGCTTGTGTCATCAACGGATA 7187

```

```

QY 1380 TGAAGAAATCAGCGGATGTCCTTGAAGAAATCTCAGAACTTCCAGTGGATCCGATC 1439
Db 7188 TGAAGAAATCAGCGGATGTCCTTGAAGAAATCTCAGAACTTCCAGTGGATCCGATC 7247
QY 1440 TGAAGTGTAAACGATGCTGTGATGAGCAGCGCATATGCCATGTTCAAGTCAAGGCGAT 1499
Db 7248 TGAAGTGTAAACGATGCTGTGATGAGCAGCGCATATGCCATGTTCAAGTCAAGGCGAT 7307
QY 1500 TCAGAGGAAAGACAAATAAGTAAAGCTTAAGACGGGATATCGAATTTATGA 1559
Db 7308 TCAGAGGAAAGACAAATAAGTAAAGCTTAAGACGGGATATCGAATTTATGA 7367
QY 1560 TAAAGTGTGTGATCAGATCTGCACTGCAATATCCAACTCAGGATTTGCTGATGCGA 1619
Db 7368 TAAAGTGTGTGATCAGATCTGCACTGCAATATCCAACTCAGGATTTGCTGATGCGA 7427
QY 1620 TACCAATATTTTTCAGGACACAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGATC 1679
Db 7428 TACCAATATTTTTCAGGACACAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGATC 7487
QY 1680 GTCAAAACCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739
Db 7488 GTCAAAACCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7547
QY 1740 TGTCTCATGAGACGGGATGCAAAACAGTGTATGCTGATGATGATGATGATGATGATGAT 1799
Db 7548 TGTCTCATGAGACGGGATGCAAAACAGTGTATGCTGATGATGATGATGATGATGATGAT 7607
QY 1800 GA 1801
Db 7608 GA 7609

```

```

RESULT 5
AAN50226
ID AAN50226 standard; DNA; 24593 BP.
XX
XX AAN50226;
AC 25-MAR-2003 (updated)
DT 16-OCT-1991 (first entry)
XX
XX DE Sequence of opine synthase gene.
XX KM Plant vector; Ti plasmid; T-DNA; ss.
XX OS Agrobacterium tumefaciens Ti plasmid.
XX PN EP140556-A.
XX PD 08-MAY-1985.
XX PF 12-SEP-1984; 84BP-0306233.
XX PR 14-SEP-1983; 83US-0532280.
XX PA (AGRK ) AGRIGENETICS RES ASSOC LTD.
XX PA (LUBR ) LUBRIZOL GENETICS INC.
XX PI Dahl GA, Sutton DW, Barker RF;
XX WPI; 1985-112088/19.
XX DR
XX PT Plasmid contg. opine synthase gene for selection - and foreign
XX PT DNA, useful as vector for transforming plant cells
XX PS Disclosure; Page 212-217; 69pp; Japanese.
XX CC Plant cells (and protoplasts) and plasmids contg. the DNA fragment
XX CC which includes an opine synthase gene plus a gene for antibiotic
XX CC resistance are claimed. These plasmids provide max. efficiency for
XX CC transfer of foreign genes and can be amplified in the plant genome.

```

CC They do not contain genes specifying tumour formation and will not
CC spread antibiotic resistance throughout the plant population.
CC (Updated on 25-MAR-2003 to correct PE field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX

SQ Sequence 24593 BP; 6554 A; 5493 C; 5787 G; 6759 T; 0 other;

Query Match 97.8%; Score 1762; DB 6; Length 24593;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1790; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

```
QY 1 ATGTGACCTTACCTCTCTTGATTAACAGTGCATATCTCCCAACCAAAATGTGGAT 60
DB 5809 ATGTGACCTTACCTCTCTTGATTAACAGTGCATATCTCCCAACCAAAATGTGGAT 5868
QY 61 CTGACAAATGTCATTAAGGCGGATGAATTGACCCGCGGTTTCCGATGCTTTAGAA 120
DB 5869 CTGACAAATGTCATTAAGGCGGATGAATTGACCCGCGGTTTCCGATGCTTTAGAA 5928
QY 121 CGAAGACCTTCTAGGGGAGAGGATTACTCAAGCTCCAACGAGTGCAGCGTGGTTA 180
DB 5929 CGAAGACCTTCTAGGGGAGAGGATTACTCAAGCTCCAACGAGTGCAGCGTGGTTA 5988
QY 181 GCTTGCAGAAAGCGTGGCGATGTCGCTTCCCGAGATCTCAGCTGTGGAAGTAGCA 240
DB 5989 GCTTGCAGAAAGCGTGGCGATGTCGCTTCCCGAGATCTCAGCTGTGGAAGTAGCA 6048
QY 241 GTTCTCTCCGCTTATATCTATTTGGCAAGAAATTTGGGCGGATCTTGAATGAAA 300
DB 6049 GTTCTCTCCGCTTATATCTATTTGGCAAGAAATTTGGGCGGATCTTGAATGAAA 6108
QY 301 CCTTGGCGCGGGCAACAGTAGTGCTGCTGGCATGCACTTGAACAATTTGCATG 360
DB 6109 CCTTGGCGCGGGCAACAGTAGTGCTGCTGGCATGCACTTGAACAATTTGCATG 6168
QY 361 GATTTTCCGAGACACAATTAATCCAAGCCCTGTTTTGCTGACGCGTAAGATGTGA 420
DB 6169 GATTTTCCGAGACACAATTAATCCAAGCCCTGTTTTGCTGACGCGTAAGATGTGA 6228
QY 421 CCGATGATTTTATGTCATTTGCGGCACTTCAATCTTAAGCTGCGGCTTTGCAAC 480
DB 6229 CCGATGATTTTATGTCATTTGCGGCACTTCAATCTTAAGCTGCGGCTTTGCAAC 6288
QY 481 CTGCAATGCGCGTGAACGAGATGCGAGATGAAGTGGTTACCGGTTTACATTAAC 540
DB 6289 CTGCAATGCGCGTGAACGAGATGCGAGATGAAGTGGTTACCGGTTTACATTAAC 6348
QY 541 CTTGAAGGGGCGTGCATTTGACATGTAAGTTAGTGAACCTGATGCTGAAGGGT 600
DB 6349 CTTGAAGGGGCGTGCATTTGACATGTAAGTTAGTGAACCTGATGCTGAAGGGT 6408
QY 601 TCGGCAAGTTCCTTCCCAACAATCGACTTCTACGACTGACAGACGTTTTTTGACAA 660
DB 6409 TCGGCAAGTTCCTTCCCAACAATCGACTTCTACGACTGACAGACGTTTTTTGACAA 6468
QY 661 TGTTCGATAGTGAAGGATCGGCTTCTTCCGGAAGATGTTCTTAAGCCGGAAGTGGG 720
DB 6469 TGTTCGATAGTGAAGGATCGGCTTCTTCCGGAAGATGTTCTTAAGCCGGAAGTGGG 6528
QY 721 GTCAATTTGGCGCTGCAATTTCCGAGACTCGTGTGGCAAAAGAACTGCTTCAAGT 780
DB 6529 GTCAATTTGGCGCTGCAATTTCCGAGACTCGTGTGGCAAAAGAACTGCTTCAAGT 6588
QY 781 GACGATGTTACAAATATATGAAGCAAGTGAATGTTGAGAGCAAGCTTTTGTCATGCT 840
DB 6589 GACGATGTTACAAATATATGAAGCAAGTGAATGTTGAGAGCAAGCTTTTGTCATGCT 6648
QY 841 TTCAGGAGGCTCCTAGTGTGCGGCGGAAATGGGGCGATGGAATTTCTCTGCTGCA 900
DB 6649 TTCAGGAGGCTCCTAGTGTGCGGCGGAAATGGGGCGATGGAATTTCTCTGCTGCA 6708
QY 901 TTCTGCTGTTTTTCTCTCTCGAGCTTACGCGCTGCTTTCGATGAGAGCGGTTCCCAAT 960
DB 901 TTCTGCTGTTTTTCTCTCTCGAGCTTACGCGCTGCTTTCGATGAGAGCGGTTCCCAAT 960
```

```
DB 6709 TTCTGCTGTTTTTCTCTCTCGAGCTTACGCGCTGCTTTCGATGAGAGCGGTTCCCAAT 6768
QY 961 CCGGACAGTGAGACCTTACTGTCCTACCAAGGCGTCCAAATGAGGAAAGCCGG 1020
DB 6769 CCGGACAGTGAGACCTTACTGTCCTACCAAGGCGTCCAAATGAGGAAAGCCGG 6828
QY 1021 CAGCTCCACCGAGCTGTTCCATCGGTTTCAACAGTGTGGCGTCTTCTTGAAGAC 1080
DB 6829 CAGCTCCACCGAGCTGTTCCATCGGTTTCAACAGTGTGGCGTCTTCTTGAAGAC 6888
QY 1081 GGTTCATAGAGAGATATGTTGCTTGGCTGCTGCTGCTATTAATCAAGGCTTGA 1140
DB 6889 GGTTCATAGAGAGATATGTTGCTTGGCTGCTGCTGCTATTAATCAAGGCTTGA 6948
QY 1141 TCAGACAAATTAAGTGGGCTCATGCTCTTGGCAAAATTTGGCTGAACGTTTGGGAG 1200
DB 6949 TCAGACAAATTAAGTGGGCTCATGCTCTTGGCAAAATTTGGCTGAACGTTTGGGAG 7008
QY 1201 GAGTCTTCTCTCAGAGGATAGAGAGATCTTCTTGGGACACATCTCTGGTGGT-AA 1259
DB 7009 GAGTCTTCTCTCAGAGGATAGAGAGATCTTCTTGGGACACATCTCTGGTGGTAA 7068
QY 1260 ACATGAGATTTTCTCATGATTTGGGACCTATTCAAGCTAATGGGAAATAGATCTGGCG 1319
DB 7069 ACATGAGATTTTCTCATGATTTGGGACCTATTCAAGCTAATGGGAAATAGATCTGGC -GG 7127
QY 1320 GTTTGGTCCAGTTTTTGAAGGGGTTTATGAGATCTCTCGCTTGGTCAACGGGAT 1379
DB 7128 GTTTGGTCCAGTTTTTGAAGGGGTTTATGAGATCTCTCGCTTGGTCAACGGGAT 7187
QY 1380 TGAAGAAATTCAGCGGATGTCCTGGAAGAAATCTCAGAACTTCCAGTCCGATGCGATC 1439
DB 7188 TGAAGAAATTCAGCGGATGTCCTGGAAGAAATCTCAGAACTTCCAGTCCGATGCGATC 7247
QY 1440 TGAAGTGTAAAGGTTGTGTGTGTGAAGCCGCAATATGCAATGTTCAAGTCAAGGCGAT 1499
DB 7248 TGAAGTGTAAAGGTTGTGTGTGTGTGAAGCCGCAATATGCAATGTTCAAGTCAAGGCGAT 7307
QY 1500 TCAGAGGAAAGACAAAATTAAGATTAAGGCTTAAAGCGGGAATCTGAACCTTTATGA 1559
DB 7308 TCAGAGGAAAGACAAAATTAAGATTAAGGCTTAAAGCGGGAATCTGAACCTTTATGA 7367
QY 1560 TAAAGTGTGTGATCAGCTGGAATGCAATTCACACTCAGGCAATGCTGACATGCGA 1619
DB 7368 TAAAGTGTGTGATCAGCTGGAATGCAATTCACACTCAGGCAATGCTGACATGCGA 7427
QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGTTGTAACAGCATATGACAGATC 1679
DB 7428 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGTTGTAACAGCATATGACAGATC 7487
QY 1680 GTCAAAATCTTCTGATGATGTAACGAAATTTCTGTTAGACATATCTCTCCGCTTG 1739
DB 7488 GTCAAAATCTTCTGATGATGTAACGAAATTTCTGTTAGACATATCTCTCCGCTTG 7547
QY 1740 TGTCCATGAGACGGGATGCAAAAGCAAGTGAATGCTGAGCTATGAGTCCAGGATCC 1799
DB 7548 TGTCCATGAGACGGGATGCAAAAGCAAGTGAATGCTGAGCTATGAGTCCAGGATCC 7607
QY 1800 GA 1801
DB 7608 GA 7609
```

RESULT 6
AAV69756
ID AAV69756 standard; DNA; 2211 BP.
XX
XX AAV69756;
XX
XX
XX 01-MAR-1999 (first entry)
XX
XX Tryptophan oxygenase iaam gene.

KM Tryptophan oxygenase; iaam gene; auxin; indoleacetic acid;
KM transgenic plant; seedless fruit; tomato; watermelon; cucumber; da.
XX
OS Agrobacterium tumefaciens.
XX
FH Key Location/Qualifiers
FT CDS 4..2205
FT /tag= a
XX
XX MO9849888-A1.
XX
XX 12-NOV-1998.
XX
PD 06-MAY-1998; 98MO-US09013.
XX
PF 06-MAY-1997; 97US-0045725.
XX
PR (UNIV) UNIV KANSAS STATE RES FOUND.
XX
PA L4 Y;
XX
PI WPI; 1999-034673/03.
XX
DR P-PSDB; AAW81574.
XX
PT A new construct to express phytohormones in developing fruit -
XX useful for, e.g. producing substantially seedless fruit from
XX transgenic plants
XX
XX Example 1; Page 21-26; 49pp; English.
XX
XX This is the nucleotide sequence of the Agrobacterium tumefaciens
XX iaam gene that codes for tryptophan oxygenase (see AAW81574). This
XX enzyme catalyses the conversion of tryptophan to indoleacetic acid
XX in the biosynthetic pathway of the auxin indoleacetic acid. A
XX claimed DNA construct comprises either an isopentenyl transferase
XX (see AAW81575) or a tryptophan oxygenase encoding sequence, operably
XX linked to an ovary or developing fruit-specific plant-expressible
XX promoter (see AAW69755 and AAW69759). The construct is used to stably
XX integrate enzymes involved in cytokinin or auxin biosynthesis into
XX the plant genome to achieve a transgenic plant (preferably tomato,
XX cucumber or watermelon) producing seedless fruit in the absence of
XX pollination.
XX
SQ Sequence 2211 BP; 528 A; 504 C; 595 G; 584 T; 0 other;
Query Match 94.2%; Score 1697.4; DB 20; Length 2211;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1726; Conservative 0; Mismatches 11; Indels 2; Gaps 2;
QY 64 ACAATGTCGATTAAGCGGATGTAATGACCGAGGTTCCGATGCTTCTTAGAAGA 123
DB 1 ACAATGTCGATTAAGCGGATGTAATGACCGAGGTTCCGATGCTTCTTAGAAGA 60
QY 124 GAAGCTTCTAGGGAGAGAGATTACTCAAGCTCCACGAGTGACGCGTGGTTAGCT 183
DB 61 GAAGCTTCTAGGGAGAGAGATTACTCAAGCTCCACGAGTGACGCGTGGTTAGCT 120
QY 184 TGCAGAAAGGCTGGCCGATGCTTCCCGAGATCTCAGCTGTGAAAGTGACGTT 243
DB 121 TGCAGAAAGGCTGGCCGATGCTTCCCGAGATCTCAGCTGTGAAAGTGACGTT 180
QY 244 CTCCTCCCTATATCTATTTGGCAAGAAATTCGGGGCGGATCTTGAATGAAACCT 303
DB 181 CTCCTCCCTATATCTATTTGGCAAGAAATTCGGGGCGGATCTTGAATGAAACCT 240
QY 304 TGGGCGGCGGCAACAGTAGTGTCTCTGTCGATCGACTTGACACACATTTGATGAT 363
DB 241 TGGGCGGCGGCAACAGTAGTGTCTCTGTCGATCGACTTGACACATTTGATGAT 300
QY 364 TTCTCCGAGACCAACTAATTCAGGCGCTTTTGTCTGAGCGGTAAAGATGTCACCG 423
DB 301 TTCTCCGAGACCAACTAATTCAGGCGCTTTTGTCTGAGCGGTAAAGATGTCACCG 360

QY 424 ATTGATCTTAGTCAATTCCTGGCCATTTCATCTCTAAGACATCCGCGCTTTCGAACCTG 483
DB 361 ATTGATCTTAGTCAATTCCTGGCCATTTCATCTCTAAGACATCCGCGCTTTCGAACCTG 420
QY 484 CCAATCCCGCTGTAGCAGAAATGACACAGTAAATGCGTTACCGGGTTTACATACCTT 543
DB 421 CCAATCCCGCTGTAGCAGAAATGACACAGTAAATGCGGGTTTACATACCTT 480
QY 544 GAAAGGCGCGTGCATTTGACATGATGATGTTATGTCGAAACCTGATGCGAAGGGTTG 603
DB 481 GAAAGGCGCGTGCATTTGACATGATGATGTTATGTCGAAACCTGATGCGAAGGGTTG 540
QY 604 GCAGTTCTCTTCCAAACAATCGACTTCTAAGACATGACACCGTTTTCACATGT 663
DB 541 GCAGTTCTCTTCCAAACAATCGACTTCTAAGACATGACACCGTTTTCACATGT 600
QY 664 TCCGATAGTGAACGATCGGCTTCTTCCGAGAGATGTTCTAAGCCGAAAGTGGCGT 723
DB 601 TCCGATAGTGAACGATCGGCTTCTTCCGAGAGATGTTCTAAGCCGAAAGTGGCGT 660
QY 724 ATTGGGCGTGGCATTTCCGGAATCGGTGTGGCAACGAACTGCTGATGCTGGGGTAGAC 783
DB 661 ATTGGGCGTGGCATTTCCGGAATCGGTGTGGCAACGAACTGCTGATGCTGGGGTAGAC 720
QY 784 GATGTTACAATATATAGAACAAAGTAGATCGTGTGAGGCAAGCTTGTGCATGCTTTC 843
DB 721 GATGTTACAATATATAGAACAAAGTAGATCGTGTGAGGCAAGCTTGTGCATGCTTTC 780
QY 844 AGGAGCGCTCTTAGTGTCTGTGCGCAATGGGGCGATGCGATTTCTCTGTCATTC 903
DB 781 AGGAGCGCTCTTAGTGTCTGTGCGCAATGGGGCGATGCGATTTCTCTGTCATTC 840
QY 904 TGCTTTCTTTCTCTCGAGGCTTACGGGCTGTCTTCGATGAGGCGTTCCCAATCC 963
DB 841 TGCTTTCTTTCTCTCGAGGCTTACGGGCTGTCTTCGATGAGGCGTTCCCAATCC 900
QY 964 GGCACAGTCGACACTTACTTGTGTCTCAAGAGCGTCAATACATGTGAAAGCGGGCAG 1023
DB 901 GGCACAGTCGACACTTACTTGTGTCTCAAGAGCGTCAATACATGTGAAAGCGGGCAG 960
QY 1024 CTGCCACCGAAGCTGTTCATCGCGTTTACACAGGTTGGCGTCTTCTTGAAGACGCT 1083
DB 961 CTGCCACCGAAGCTGTTCATCGCGTTTACACAGGTTGGCGTCTTCTTGAAGACGCT 1020
QY 1084 TTTGATGAGGAGATATGTTGTTGGCTGCGCTGTGCTATTACTAGGCGCTTGAATCA 1143
DB 1021 TTTGATGAGGAGATATGTTGTTGGCTGCGCTGTGCTATTACTAGGCGCTTGAATCA 1080
QY 1144 GGAACATTAAGTGGGCTCATGACTCTGCGCAATTTGGCTGAACCGTTTCGGAGGGAG 1203
DB 1081 GGAACATTAAGTGGGCTCATGACTCTGCGCAATTTGGCTGAACCGTTTCGGAGGGAG 1140
QY 1204 TCTTTCTCTTACAGGATTAAGAGGATCTTCTGGGACACATCTCTCTGTGTGT -AACA 1262
DB 1141 TCTTTCTCTTACAGGATTAAGAGGATCTTCTGGGACACATCTCTCTGTGTGT -AACA 1200
QY 1263 TGAAGTTTCTCTATATGAGGACTTTCCTGAGGACATCTCTGCTGTGTGAAACA 1322
DB 1201 TGAAGTTTCTCTATATGAGGACTTTCCTGAGGACATCTCTGCTGTGTGAAACA 1259
QY 1323 TGTGTCAGTTTGAAGAGGGGTTTATGAGATCTCTCGCTTGTGATCAACGGAATGA 1382
DB 1260 TGTGTCAGTTTGAAGAGGGGTTTATGAGATCTCTCGCTTGTGATCAACGGAATGA 1319
QY 1383 AGAAATACAGCGATGTGCTTGAAGGAATCTAGAACTTCCAGCTGCGATGCGATCTGA 1442
DB 1320 AGAAATACAGCGATGTGCTTGAAGGAATCTAGAACTTCCAGCTGCGATGCGATCTGA 1379
QY 1443 AGTGTGTAACGAGTGTCTGTGAGCAGGCGCATATGTCATGTTCAAGTCAAGGCGATTCA 1502
DB 1380 AGTGTGTAACGAGTGTCTGTGAGCAGGCGCATATGTCATGTTCAAGTCAAGGCGATTCA 1439
QY 1503 GAAGGAAAGACAAAATAAAGATTAAGGCTTAAAGCGGGATATCTGAATCTTATGATTA 1562

DB 1440 GAAGGAAAAACAAAATAAGATTAAGCGGATATCTGAACCTTATGATTA 1499
QY 1563 GGTGGGTGTCACATCTGGACTCCGAAATATCCAACTAGCACTTGGCATGCCATAC 1622
DB 1500 GGTGGGTGTCACATCTGGACTCCGAAATATCCAACTAGCACTTGGCATGCCATAC 1559
QY 1623 CAATATTTTTCAGGACCACTGAAACCAAGCGGTGATTAAGCCATATGACAGATGTC 1682
DB 1560 CAATATTTTTCAGGACCACTGAAACCAAGCGGTGATTAAGCCATATGACAGATGTC 1619
QY 1683 AAAAATCTTCTGATGATGATGACGAAATTTCTGGTTAGACATATCTCCGCTTGTGT 1742
DB 1620 AAAAATCTTCTGATGATGATGACGAAATTTCTGGTTAGACATATCTCCGCTTGTGT 1679
QY 1743 CCTCATGAGAGGATGCGAAAGACGATGTTGCTTGGACTATGATGATGATGATGATGAT 1801
DB 1680 CCTCATGAGAGGATGCGAAAGACGATGTTGCTTGGACTATGATGATGATGATGATGATGAT 1738

RESULT 7

AAT48868/C
ID AAT48868 standard; DNA; 6112 BP.

AC AAT48868;
DT 07-APR-1997 (first entry)

DE Iams expression vector pPH5838.

KW Male sterile plant; hybrid seed; pollen; microspore; oilseed;

KM canola; rape; antisense RNA; stress resistance;

KM herbicide resistance; transgenic plant; crop protection;

KM Iams; indole acetamide synthase; Bp10; promoter; Brassica;

KM pPH5838; ds; cycl1c.

OS Chimeric Brassica sp.;

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

OS Chimeric Agrobacterium tumefaciens.

CC from a male fertile line.
XX
SQ Sequence 6112 BP; 1515 A; 1466 C; 1451 G; 1679 T; 1 other;
Query Match 88.9%; Score 1600.4; DB 18; Length 6112;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 1689; Conservative 0; Mismatches 111; Indels 2; Gaps 2;

QY 1 ATGTGACCTTCACTCTCTCTGATTAACAGTGGATCATCTCCCAACCAAAATGATGAT 60
DB 2870 ATGTGACCTTCACTCTCTCTGATTAACAGTGGATCATCTCTCCCAACCAAAATGATGAT 2811
QY 61 CTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 2810 CTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2751
QY 121 CGAAGACCTTCTAGGGAAGAGATTAATCAAGTCTCAACCGAGTGCAGCGTGGTTA 180
DB 2750 CGTGAAGCTTCTAGGGAAGAGATTAATCAAGTCTCAACCGAGTGCAGCGTGGTTA 2691
QY 181 GCTTGCAGAAAGCTGGCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 2690 GCTTGCAGAAAGCTGGCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 2631
QY 241 GTTCTCTCGCTATATCTATATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 2630 GCTTCTCTCGCTATATCTATATGATGATGATGATGATGATGATGATGATGATGATGAT 2571
QY 301 CCTTGGGCGCGGCAACAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 2570 CCTTGGGCGCGGCAACAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 2511
QY 361 GATTTCTCGGACACACATTAATCAAGCTGTTTGTGTCGAGCGGTAAGATGATGAT 420
DB 2510 GATTTCTCGGACACACATTAATCAAGCTGTTTGTGTCGAGCGGTAAGATGATGATGAT 2451
QY 421 CGATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 2450 TCAGGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2391
QY 481 CTGCAATGCGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 2390 CTGCAATGCGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2331
QY 541 CTGAAAGGCGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 2330 CTGAAAGGCGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2271
QY 601 TCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 2270 TCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2211
QY 661 TGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 2210 TGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2151
QY 721 GTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 2150 ATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2091
QY 781 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 2090 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2031
QY 841 TTCAGGAGCTCTCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 2030 TTCAGGAGCTCTCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1971
QY 901 TTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1970 TTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1911
QY 961 CCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020

QY 361 GATTTCGAGACACAACTAATCCAGCCCTGTTTTCGTGACGGTAAAGATGCA 420
Db 2564 GATTTCGAGACACAACTTCTCAACCTGTTTTCGTGACGGTAAAGATGCA 2505
QY 421 CCGATTGATTTAGTCACTTTGCGCCATTTCATCTTAAGA CTGCGGCTTTTCAACC 480
Db 2504 TCAGAGGATCTTGTGATTTGCGCCATTTCATCTTAAGA CTGCGGCTTTTCAACC 2445
QY 481 CTGCGCAATGCGGCTTAAGCAAGATGAGCAAGATGAATGCGTTACCGGTTTACATPACC 540
Db 2444 CTGCAATATGCGGCTTAAGCAAGATGAGCAAGATGAATGCGTTACCGGTTTACCTGACC 2385
QY 541 CTGGAAGGGGCGGCTGCACTTTGACATGATGCTTAAGTGAACCTGATGCTGAAGGCT 600
Db 2384 CTGGAAGGGGCGGCTGCACTTTGACATGATGCTTAAGTGAACCTGATGCTGAAGGCT 2325
QY 601 TCGGCAAGTTCTTTTCAACAAATCGAATTGCTTACGACTGACAGCCGTTTTCACAA 660
Db 2324 TCGGCAAGTTCTTTTCAACAAATGACTTCTTATGACTACAGATGCTTTTTCACAA 2265
QY 661 TGTTCGATGATGAGGAGATGCGGCTTTTCCGAGAGATTTCTCTAAGCCGAAGTGGCG 720
Db 2264 TGTTCGATGATGAGGAGATGCGGCTTTTCCGAGAGATTTCTCTAAGCCGAAGTGGCG 2205
QY 721 GTCAATTGCGGCTGCACTTTCCGAGCTGATGCTGGAACCAAGCTGCTTCACTGAGGATA 780
Db 2204 ATCATTTGGGCTGCACTTTCCGAGCTGATGCTGGAACCAAGCTGCTTCACTGAGGATA 2145
QY 781 GACGATGTTAACAATATATGAGCAAGTGAATGATGCTGATGAGCAAGCTTTGCTCACAATGCT 840
Db 2144 GACGATGTTAACAATATATGAGCAAGTGAATGATGCTGATGAGCAAGCTTTGCTCACAATGCT 2085
QY 841 TTCAGGAGCGCTCTCTAGTGTCTGCGCCGAATGCGGCGCATGCTGCTCTGCTGCA 900
Db 2084 TTCAGGAGCGCTCTCTAGTGTCTGCGCCGAATGCGGCGCATGCTGCTCTGCTGCA 2025
QY 901 TTCTGCTGTTTCTTCTCTGAGCGTTACGCGCTGCTTCAATGAGCGGCTTCCCAAT 960
Db 2024 TTCTGCTGTTTCTTCTCTGAGCGTTACGCGCTGCTTCAATGAGCGGCTTCCCAAT 1965
QY 961 CCGGCGACAGTCGACACTTACTTGTCTACCAAGGCGTCCCAATACATGTTGAAGCGCGG 1020
Db 1964 CCGGCGACAGTCGACACTTACTTGTCTACCAAGGCGTCCCAATACATGTTGAAGCGCGG 1905
QY 1021 CAGCTGCCACCGAAGCTTTCCATCGGCTTTACAAAGGTTGGCGGCTTTTGAAGAAC 1080
Db 1904 CAGCAAGCGACCGAAGCTTTCCATCGGCTTTACAAAGGTTGGCGGCTTTTGAAGAAC 1845
QY 1081 GGTTCATGAGGAGATATGTTGTTGGCTTGGCTGTGCTATTACTCAGGCGCTTGA 1140
Db 1844 GGTTCATGAGGAGATATGTTGTTGGCTTGGCTGTGCTATTACTCAGGCGCTTGA 1785
QY 1141 TCAGGACATTAAGTGGGCTCATGACTCTGCGCAATTTGCGTGAACGTTTGGGAGG 1200
Db 1784 TCAGGACATTAAGTGGGCTCATGACTCTGCGCAATTTGCGTGAACGTTTGGGAGG 1725
QY 1201 GAGTCTTCTCTTCAAGGATGAGAGATTTTCTGGGCAACATCTCTCTGTGTGT-AA 1259
Db 1724 GAGTCTTCTCTTCAAGGATGAGAGATTTTCTGGGCAACATCTCTCTGTGTGTAA 1665
QY 1260 ACATGAGTTTCTCATGATTTGGGACCTATTCAACTAATGGAATPAGGATGCGCGG 1319
Db 1664 ACATGAGTTTCTCATGATTTGGGACCTATTCAACTAATGGAATPAGGATGCGC-GG 1606
QY 1320 GTTTCGTCAGTTTGAAGCGGTTTATGAGATCTCTGCTGATCAACCGATA 1379
Db 1605 GTTTCGTCAGTTTGAAGCGGTTTATGAGATCTCTGCTGATCAACCGATA 1546
QY 1380 TGAAGAAATTCAGCGGATGCTCTGAAGGAATCTCAAACTTCCAGCTGCGATCCATC 1439
Db 1545 TGAAGAAATTCAGCGGATGCTCTGAAGGAATCTCAAACTTCCAGCTGCGATCCAC 1486
QY 1440 TGAAGTGTAAACGGTGTCTGTGAGCAAGCGCATATGCAATGTTCAAGTCAAGCGCAT 1499

Db 1485 TCAGTGTGTTAAGGTGTGTCTGTAAAGCCAGCTATACGCAATGTTCAATCAGGCGAT 1426
QY 1500 TCAGAGAAAGAAACAAAATTAAGATTAAGCTTTAAGCGGAGATCTGAACCTTAATGA 1559
Db 1425 TGAGAAAGAAAGAAACAAAATTAAGATTAAGCTTTAAGCGGAGATCTGAACCTTAATGA 1366
QY 1560 TAAGTGTGTGTACATCTGAGTCTGCAATATCCAACTCAGGCAATTCCTGACATGCCA 1619
Db 1365 TAAGTGTGTGTACATCTGAGTCTGCAATATCCAACTCAGGCAATTCCTGACATGCCA 1306
QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACAGCATATGACAGATC 1679
Db 1305 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACAGCATATGACAGATC 1246
QY 1680 GTCAAAATCTTCTCTGATGATGACGAAATTTCTGTTAAGACATATCTCTCCGTTTG 1739
Db 1245 GTCAAAATCTTCTCTGATGATGACGAAATTTCTGTTAAGACATATCTCTCCGTTTG 1186
QY 1740 TGTCTCATGAGAGGATGCGAAATGAGTATGCTGAGTATGAGTGGAGATCC 1799
Db 1185 TGTCTCATGAGAGGATGCGAAATGAGTATGCTGAGTATGAGTGGAGATCC 1126
QY 1800 GA 1801
Db 1125 GA 1124

RESULT 9

AAN60970
ID AAN60970 standard; DNA; 21126 BP.

AC AAN60970;

DT 25-MAR-2003 (updated)

DT 28-OCT-1991 (first entry)

DE TL-DNA region of A. rhizogene agropine-type plasmid pRiR1.

KW Ri plasmid; plant plasmid.

OS Agrobacterium rhizogenes HRI and A4.

PN BP204590-A.

PD 10-DEC-1986.

PF 21-APR-1986; 86EP-0400855.

PR 22-APR-1985; 85US-0725368.

PA (INRG) INRA INST NAT RECH AGRONOMIQUE.

PI Slightom JL, Tepfer DA;

DR WPI, 1986-327378/50.

XX Plant cell genetic modification - using pRi T-DNA promoter to
PT control expression of heterologous foreign structure genes, e.g.
PT for phaseolin.

PS Disclosure; Fig 2; 86pp; English.

XX The TL-DNA sequence is that of an Agrobacterium Ri plasmid, it
CC contains sixteen open reading frames, bounded by eukaryotic
CC promoters, ribosome binding and polyadenylation site. The plasmid
CC sequence may be used for genetic manipulation of plants, conveying
CC phenotypes such as improved resistance to environmental stress, disease
CC and parasitism.
CC (Updated on 25-MAR-2003 to correct PA field.)

CC Sequence 21126 BP; 5359 A; 4901 C; 5308 G; 5558 T; 0 other;

SQ

Query Match 6.8%; Score 121.6; DB 7; Length 21126;
 Best Local Similarity 45.7%; Pred. No. 5,7e-28;
 Matches 540; Conservative 0; Mismatches 634; Indels 8; Gaps 3;

QY 623 TCGACTTGTCTACGACCTGACGACCGTTTGTACCAATGTTCCGATGAGCGATCG 682
 DB 7273 TTGACCTGCGGTACGACTATCAAGTTTCTCCAGAGGTGCGATGCGCATGAGGTGTGG 7332
 QY 683 GCTTCTTCCGAGAGTGTCTTAAGCCGAAAGTGGCGGTGATGGCGGTGATTTCCG 742
 DB 7333 GGCATTATCCGAGAGGACCACTACATGATGTGCACTGAGAGGCGGCTGTCTG 7392
 QY 743 GACTCGTGTGCGAAAGCACTGCTTATGCTGGGTAGACGATGTTACATATATGAAG 802
 DB 7393 GCTTGTGTGCGCAAGAACTATGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 7452
 QY 803 CAAATGATGCTGT--TGAGGCAAGCTTGTGCACTGCTTTCAGGAGCGTCTTAATG 859
 DB 7453 CCGTGTATGAGATCCGTAAGTTTGGGGCATCGCCGATGCCAAAGCGGCGATCACAGG 7512
 QY 860 TCGTGGCCGAAATGGGGCGCATGCGATTTCTCTGCTGATTTCTGCTTTTCTTCC 919
 DB 7513 CCTTGAAGTGTGCGGTGCTGCTTCTTCTCCGCAACAACTTGTGCTGCACTATC 7572
 QY 920 TCGAGCGTTACGCGCTGTCTTTCATGAGCGCGTTCCTCCAAATCCCGGACAGTGCACCT 979
 DB 7573 TGGATTAAGTTTAATATCCGTCCAGCGCTTGTCTTCTTGTGCGGAAAGACACACAG 7632
 QY 980 ACTTGTCTACCAAGCGGTCAATATCATGTGAAAGCGGCGGACGTCGCCAAGAGCTGT 1039
 DB 7633 CACTATATTTCCCGCAGAAAGCTACGATGCGACGCGGGGCAAGCTCCCGGGGATAT 7692
 QY 1040 TCCATGCGCTTACACGCTTGGCGGTGCTTCTTGAAGACGCTTTCATGACGAGATA 1099
 DB 7693 TTGAGCGGTGACATGTGCGATGGAAGACATCTACTACCAAGGTGTGAACGGAATGCA 7752
 QY 1100 TTGTGTGCTTGCCTGCTGCTATTTACTCAGGCGCTTGAATACGACACATTAAGTGG 1159
 DB 7753 GAGAGCTGATGCTCCGATGATATCTTTCATGTTGAAGACGCTGCTGATGAG 7812
 QY 1160 CTGATGACTCTGCAAAATTTGGCTGAACGTTTCCGAGAGGAGTCTTCTTTCAGGGA 1219
 DB 7813 CCTCAGAAAGACGCGACCTTTGGCTCCGAGAGTTCCGAAATTCATCTTCCATGCGCTT 7872
 QY 1220 TAAGAGAGATCTTCTGAGGACATCTCTCTGTGTGT---AAACTGAGATTTCTC 1275
 DB 7873 TGTGTGAGATCTTCAAGCTGTGTGTAATTCAGTCTCTGTGCGAAGGCAATGCCAAC 7932
 QY 1276 ATGATGGGACCTATTTCAAGCTAATGGGAATAGGATCTGGCGGGTGTGCTCCAGTTT 1335
 DB 7933 ATGATTTGAGGCTTTCGGATCTGAGTGTGGATACGCGCGAGTTTCTCTC-TATTAC 7991
 QY 1336 GAAAGCGGTGTTATGAGATCTTCCGCTGTGCTATCAAGGATATGAAGAAATCAGCGG 1395
 DB 7992 AAGGTGTGTTTTCACGATCTGAGATGATTAATCAATGAGTACGAGGACGACAT 8051
 QY 1396 ATGTGCTTGAAGAAATCTCAAGACTTCCAGCTGCGATTCGCATCTGAAGTGTAAAGGT 1455
 DB 8052 CTTTCTAATGTGTGGGTTCAACTTTTGCAGGCTCTGTATGCGCATGAAATATTCAGAAA 8111
 QY 1456 GTGTCTGTAGCGAGCGCATATGCGCATGTTCAAGTCAAGGCGCATTCGAAGAAAGACA 1515
 DB 8112 AGCATGCGAAAGACGACTGTGTTTGTATCCGTGTGATGAGTACGCAAGAGGCGGG 8171
 QY 1516 AAAATTAAGTAAAGCTTAAAGCGGATATCTGAACCTTATATGAATAGGTGTGTACA 1575
 DB 8172 AGATTAAGGATAGCTTGAACACGCTCATTCGGGTGTTTGAACGAGTATCATTTGGC 8231
 QY 1576 TCTGAGTCCGAATATTCACATGAGCATTTGCTGACATGCGATACCAATATTTTCA 1635
 DB 8232 GCGAGTGTCTAGGCGGCTACAGTTGATTAACAGATGCGCGGAGTGAAGATCTTCTTAC 8291
 QY 1636 GCACCAATGAACCAAGGCTTGAATACAGCATATGACAGGATGCTCAAAACTCTTCTG 1695

DB 8292 TACAATATGAACCCGCGTGGAAACTGTCTGCGCTGTCAATTCAGACATCTTCATG 8351
 QY 1696 ATGACGTGAACGAAATTTGTTAGCCATATCCCTGTTGTCTCTCAATGACGCGG 1755
 DB 8352 GTACAGAACCAAAAGTTTGGTTAACTCCGCGATCCAGCATGATATGACCGAATGG 8411
 QY 1756 ATGCAAAAGCAGTGTATTCCTGAGCTATGATGTCGACGAT 1797
 DB 8412 CTTGTCCGTGAGCTGTGTTGATTCATGATGATGCGCAGCT 8453

RESULT 10
 ABL13666
 ID ABL13666 standard; cDNA; 4734 BP.
 AC ABL13666;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 35480.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ss.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EM;
 XX
 DR WPI; 2001-656860/75.
 DR P-PSDB; ABB69563.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 35480; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (ABBS7237-ABBS7072).
 CC
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 XX Sequence 4734 BP; 1394 A; 978 C; 1048 G; 1314 T; 0 other;

Query Match 2.2%; Score 39.8; DB 23; Length 4734;
 Best Local Similarity 52.0%; Pred. No. 0.19;
 Matches 89; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 690 TCGGAGATGTTCTTAAGCGCAAGTGGCGGTGATGGCGCTGCTGCTGCTGCTGCTGCTG 749
 DB 1219 TTAAGGAATATATCAACCTTAAGATCGTATGATTTGAGCTGGAATGCTGCTGCTATC 1278
 QY 750 GGTGGCAACGAACTGCTTCAATGCTGGGTGAGACGATTTCAATATATGACCAAGTGA 809

Db 1279 GGCTGCAATCACCCTTTGCAAAACGGCTGTGACGATTTCTTATTCGTGAGGACGCGGG 1338
 QY 810 TCGGTGAGGAGCAAGCTTTGTCACATGCTTTACGAGGACGCTCTAGTGT 860
 Db 1339 CCGGGTGGGCGACGACATTTGTTCATACCGCTCAGCAATATATGAGAAGT 1389

RESULT 11

ABL15662
 ID ABL15662 standard; cDNA; 4800 BP.

AC ABL15662;
 XX

DT 26-MAR-2002 (first entry)
 XX

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 41468.

KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ss.

OS Drosophila melanogaster.
 XX

PN WO200171042-A2.
 XX

PD 27-SEP-2001.
 XX

PF 23-MAR-2001; 2001WO-US09231.
 XX

PR 23-MAR-2000; 2000US-191637P.
 XX

PR 11-JUL-2000; 2000US-0614150.
 XX

PA (PEKE) PE CORP NY.
 XX

PI Venter JC, Adams M, Li PMD, Myers EW,
 XX

DR WPI; 2001-656860/75.
 XX

DR P-PSDB; ABB71559.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions -

PS Claim 1; SEQ ID NO 41468; 21bp + Sequence Listing; English.
 XX

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC
 XX

Sequence 4800 BP; 1265 A; 1116 C; 1141 G; 1278 T; 0 other;

Query Match 2.2%; Score 39.8; DB 23; Length 4800;
 Best Local Similarity 52.0%; Pred. No. 0.2;

Matches 89; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 690 TCGGAGATGTTCTTAAGCCGAAAGTGGCGGTGATTTGGCGCTGCTTCCGACTCGT 749
 Db 4496 TTCAGGAAATATATACCAAGTTCGATCATTTGAGCTGGAATGGCTGGCTTATC 4555

QY 750 GGTGGAACGAAGCTCTTCATGCTGGGGTGAAGCATGTATCAATATATGAACAAGTA 809
 Db 4556 GGTGCAAAATCACTTTTGAACGCGCTGTGACGATTTCTTATCTGAGGACGCGG 4615

QY 810 TCGTGTGAGGCAAGCTTTGTCACATGCTTTCAAGGACGCTCTAGTGT 860
 Db 4616 CCGGGTGGGCGACGACATTTGTTCATACCGCTCAGCAATATATCAGAAGT 4666

RESULT 12
 ABL13667
 ID ABL13667 standard; cDNA; 1872 BP.

AC ABL13667;
 XX

DT 26-MAR-2002 (first entry)
 XX

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 35483.

KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ss.

OS Drosophila melanogaster.
 XX

PN WO200171042-A2.
 XX

PD 27-SEP-2001.
 XX

PF 23-MAR-2001; 2001WO-US09231.
 XX

PR 23-MAR-2000; 2000US-191637P.
 XX

PR 11-JUL-2000; 2000US-0614150.
 XX

PA (PEKE) PE CORP NY.
 XX

PI Venter JC, Adams M, Li PMD, Myers EW,
 XX

DR WPI; 2001-656860/75.
 XX

DR P-PSDB; ABB69564.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions -

PS Claim 1; SEQ ID NO 35483; 21bp + Sequence Listing; English.
 XX

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC
 XX

Sequence 1872 BP; 499 A; 477 C; 524 G; 372 T; 0 other;

Query Match 2.2%; Score 39.2; DB 23; Length 1872;
 Best Local Similarity 52.4%; Pred. No. 0.17;

Matches 86; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 690 TCGGAGATGTTCTTAAGCCGAAAGTGGCGGTGATTTGGCGCTGCTTCCGACTCGT 749
 Db 219 TTCAGGAAATATATACCAAGTTCGATCATTTGAGCTGGAATGGCTGGCTTATC 278

QY 750 GGTGGAACGAAGCTCTTCATGCTGGGGTGAAGCATGTATCAATATATGAACAAGTA 809
 Db 279 GGTGCAAAATCACTTTTGAACGCGCTGTGACGATTTCTTATCTGAGGACGCGG 338

QY 810 TCGTGTGAGGCAAGCTTTGTCACATGCTTTCAAGGACGCTC 853
 Db 339 CCGGGTGGGCGACGACATTTGTTCATACCGCTCAGCAATATATC 382

RESULT 13
 AAX15122
 ID AAX15122 standard; cDNA to mRNA; 1575 BP.

XX
AC AAX15122;
XX
DT 23-APR-1999 (first entry)
XX
DE Nucleic acid encoding an apoptosis inducing protein.
XX
KM Cell death; apoptosis; inhibition; proliferation; cancer cell;
XX apoptosis inducing protein; AIP; chub mackerel; anticancer; ss.
XX
OS Scomber japonicus.
XX
FH Key Location/Qualifiers
FT CDS 1..1575
FT /*tag= a
XX
PN MO9852972-A1.
XX
PD 26-NOV-1998.
XX
PF 22-MAY-1998; 98MO-JP02261.
XX
PR 23-MAY-1997; 97JP-0133549.
XX
PA (TENS-) TENSEI SUISAN CO LTD.
XX
PI Iwamoto M, Jung S;
XX
PI WPI; 1999-070139/06.
XX
DR P-PSDB; AAW96805.
XX
PT Apoptosis inducing protein from mackerel and gene encoding it - for
PT use as anticancer agents
XX
PS Claim 5; Page 32-34; 47pp; Japanese.
XX
CC The present sequence encodes a protein which induces cell death
CC (apoptosis) and inhibits the proliferation of cancer cells. The
CC protein (apoptosis inducing protein, AIP) is isolated from chub
CC mackerel. The protein can be used as an anticancer agent and as a
CC reagent for study of the mechanisms of apoptosis in vitro.
XX
SQ Sequence 1575 BP; 436 A; 358 C; 406 G; 375 T; 0 other;
XX
Query Match 2.1%; Score 38.6; DB 20; Length 1575;
Best Local Similarity 66.1%; Pred. No. 0.25; Mismatches 34; Indels 3; Gaps 1;
Matches 72; Conservative 0;
XX
QY 713 AAGTGGCGATCATTTGGCGCTGGCATTTCCGAGCTCGTGTGGCAAGCACTGCTTCATG 772
DB 182 ATGTGTTATAGTCGAGCTGGCATGGCCGAGCTGACGGCGCCAACTTCTGCAAGACG 241
XX
QY 773 CTGGGGTAGACGATGTATCAATATATGAGCAAGTGAATCGTTGGAGG 821
DB 242 CAGG---ACACACGGTAACCATATGTGAGGCTAATGATCGTTGGAGG 287
XX
RESULT 14
AB256743
ID AB256743 standard; cDNA; 468 BP.
XX
AC AB256743;
XX
DT 28-MAR-2003 (first entry)
XX
DE Aspergillus oryzae polynucleotide SEQ ID NO 5856.
XX
KM Aspergillus oryzae; fermentation; fungus; industrial; EST;
XX expressed sequence tag; gene; ss.
XX
OS Aspergillus oryzae.
XX
PN WO200279476-A1.

XX
PD 10-OCT-2002.
XX
XX 22-MAR-2002; 2002MO-IB00890.
XX
PR 30-MAR-2001; 2001JP-0098371.
XX
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX (NAPR-) NAT RES INST BREWING.
XX (NORQ) NAT FOOD RES INST MIN AGRIC.
XX
PI Machida M, Akita O, Kaehiwagi Y, Kitamoto K, Horiuchi H;
PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
XX
DR WPI; 2003-046817/04.
XX
PT Detection of expression of specific Aspergillus genes for monitoring
PT the fermentation and growth conditions of the fungus, using DNA probes
XX
PS Claim 1; SEQ ID NO 5856; 48pp + Sequence Listing; Japanese.
XX
CC The invention relates to a polynucleotide having any of 6006 specific
CC sequences (AB250888-AB256893), which are expressed by a fungus under
CC specific culture conditions including one or more of eutrophic,
CC oligotrophic, solid, early germination, alkaline, high temperature, low
CC temperature or maltose culture or polynucleotides stringently hybridising
CC to these sequences. The polynucleotides are useful for monitoring the
CC progress of fermentation and the growth conditions of a fungus,
CC especially of Aspergillus oryzae which is widely used in industrial
CC fermentation. Also monitoring for fungal contamination.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 468 BP; 124 A; 141 C; 111 G; 90 T; 2 other;
XX
Query Match 2.1%; Score 38.4; DB 25; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.14; Mismatches 82; Indels 3; Gaps 1;
Matches 100; Conservative 1;
XX
QY 706 AAGCCGAAGTGGCGGTCAATTGGCGCTGGCATTTCCGAGCTCGTGTGGCAAGCACTG 765
DB 164 AACACCAATGTCGATTCGAGCTGGCGCGGATCTCAGGCTCACAGCCGCTCTCTC 223
XX
QY 766 CTTCATGCTGGGTAGACGATGTTACATATATGAAGCAAGTGAATGTTGAGGCAAG 825
DB 224 GATAGCGTGGGAATCCACAATTGGGAATATCTGAAAGCGATGACCGGTAAGAGGCGGT 283
XX
QY 826 CTTTGGTCAATGCTTTCAAGGAGCGCTCTAGTGTGTCGCGCAAAATGGGGGAGATGCGA 885
DB 284 TTCCGAACCAACTCTGTAAGGAACAG---GCGAATGGGCGCAATATGGGCGCCATGCGG 340
XX
QY 886 TTTTCT 891
DB 341 CTGCTT 346
XX
RESULT 15
AAC05199
ID AAC05199 standard; cDNA; 273 BP.
XX
AC AAC05199;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 9274.
XX
KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX

PN EPI033401-A2.
XX
XX 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 9274; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
XX SQ Sequence 273 BP; 53 A; 73 C; 66 G; 80 T; 1 other;
XX
XX Query Match 2.1%; Score 37.8; DB 21; Length 273;
XX Best Local Similarity 50.8%; Pred. No. 0.16; Mismatches 0;
XX Matches 90; Conservative 0; Indels 0; Gaps 0;

QY 867 CGAATGCGGCGCGATGCTTCTCTGCGCATTCGTTGTTTCTTCTCGAGCG 926
DB 97 CGATCGGGCTCGATCGACCTCTGCTCTCAGGTTCAAGCATTTTCTGCTCGGCGAG 156
QY 927 TTACGGCTCTTTCATGAGCGCGTTCCCAATCCCGGACAGTCGACACTTACTTGGT 986
DB 157 AGACGGGGTTTCCACATGTTGGCCAGGCTGCTCGAACCCCTGACCTCAAGTATCAGC 216
QY 987 CTACCAAGCGCTTCATATCATGTGMAAGCGGCGAGCTGCCACGAGTGTTCGA 1043
DB 217 CCACCTCGGCTTCCCAAGTGTGAGATTACAGAGTGTGATCCACTGACCCGGGTCCA 273

Search completed: November 23, 2003, 21:01:45
Job time : 483.254 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 21:01:51 / Search time 3836.59 Seconds
(without alignments)
11409.196 Million cell updates/sec

Title: US-09-434-837-10_COPY_1_1801

Perfect score: 1801
Sequence: 1 atgcagcttcacccctccct.....ctatgagtcgagatccga 1801

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estbda:*
2: em_estbhm:*
3: em_estbn:*
4: em_estbm:*
5: em_estbv:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estc1:*
10: gb_estc2:*
11: gb_estc3:*
12: gb_estc4:*
13: gb_estc5:*
14: gb_estc6:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrc:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rnd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.8	2.8	618	9	AU239500 AU239500
2	44.8	2.5	564	12	BJ123611 BJ123611
3	44.2	2.5	481	12	BJ101485 BJ101485
4	44.2	2.5	486	12	BJ103567 BJ103567

5	44.2	2.5	508	9	AU199996 AU199996
6	44.2	2.5	511	12	BJ108542 BJ108542
7	44.2	2.5	513	9	AU203802 AU203802
8	44.2	2.5	529	12	BJ120847 BJ120847
9	44.2	2.5	540	9	AU201678 AU201678
10	44.2	2.5	560	14	CB401025 CB401025
11	44.2	2.5	584	12	BJ121613 BJ121613
12	44.2	2.5	586	12	BJ121635 BJ121635
13	44.2	2.5	597	9	AU207635 AU207635
14	44.2	2.5	603	12	BJ101145 BJ101145
15	44.2	2.5	608	9	AU204573 AU204573
16	44.2	2.5	619	12	BJ117235 BJ117235
17	44.2	2.5	634	12	BJ124022 BJ124022
18	44.2	2.5	655	12	BJ103250 BJ103250
19	43.8	2.4	1201	13	BJ381961 BJ381961
20	43.4	2.4	482	12	BJ101110 BJ101110
21	43	2.4	338	9	AU109666 AU109666
22	42.6	2.4	631	12	BJ117453 BJ117453
23	42.4	2.4	850	13	BU348455 BU348455
24	42	2.3	362	14	CA926583 CA926583
25	42	2.3	1201	13	BJ376097 BJ376097
26	40.8	2.3	999	29	CNS0601H CNS0601H
27	40.6	2.3	310	13	BU821300 BU821300
28	40.6	2.3	411	9	A1102326 A1102326
29	40.6	2.3	479	14	CA934750 CA934750
30	40.6	2.3	1376	29	AG176243 AG176243
31	39.4	2.2	284	9	A1101650 A1101650
32	39.4	2.2	456	9	AA943462 AA943462
33	39.4	2.2	462	12	BJ100778 BJ100778
34	39.4	2.2	511	10	BG378762 BG378762
35	39.4	2.2	560	12	BJ296551 BJ296551
36	39.2	2.2	404	12	BJ305833 BJ305833
37	39.2	2.2	410	12	BJ11075 BJ11075
38	39.2	2.2	437	13	BY467572 BY467572
39	39.2	2.2	501	14	CA614930 CA614930
40	39.2	2.2	502	12	BJ368223 BJ368223
41	39.2	2.2	554	9	A1457056 A1457056
42	39.2	2.2	564	12	BI566474 BI566474
43	39.2	2.2	622	12	BI627861 BI627861
44	39.2	2.2	623	12	BI588865 BI588865
45	39.2	2.2	657	9	AA440930 AA440930

ALIGNMENTS

RESULT 1
LOCUS AU239500
DEFINITION AU239500 RAF19 Arabidopsis thaliana CDNA clone RAF19-89-D06 5',
618 bp mRNA linear EST 01-APR-2002
ACCESSION AU239500
VERSION AU239500.1 GI:19878669
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Aktiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
Itoh,M., Ishii,Y., Arikawa,T., Shibusaki,K., Shinozaki,K.,
M., Hayashizaki,Y., and Shinozaki,K.
Large scale analysis of Arabidopsis full-length CDNA
unpublished

TITLE JOURNAL COMMENT

Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambda phage-1 vector (Garnier et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.genec.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

source

location/Qualifiers

1..618
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL19-89-D06"
/issue_type="mixture of silique and flower"
/lab_host="DH10B"
/clone_lib="RAFL19"
/note="Site_1: BamHI; Site_2: SalI; Subtraction Library"
BASE COUNT 158 a 143 c 148 g 166 t 3 others
ORIGIN

Query Match 2.8%; Score 49.8; DB 9; Length 618;
Best Local Similarity 60.9%; Pred. No. 0.0069;
Matches 81; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

LOCUS B123611 564 bp mRNA linear EST 23-JAN-2002
DEFINITION B123611 unpublished oligo-capped cDNA library, C. elegans L1 stage
ACCESSION B123611
VERSION B123611.1 GI:18283750
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Pelodierinae; Caenorhabditis.
1 (bases 1 to 564)
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.
A complementary view of the C. elegans genome
Unpublished
Contact: Tadashi Shin-i
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: shin@genes.nig.ac.jp.
Location/Qualifiers
1..564
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1302a06"
/sex="hermaphrodite"
/issue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

FEATURES

source

1..564
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1302a06"
/sex="hermaphrodite"
/issue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

BASE COUNT 170 a 123 c 115 g 151 t 5 others
ORIGIN

Query Match 2.5%; Score 44.8; DB 12; Length 564;
Best Local Similarity 56.6%; Pred. No. 0.19;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

LOCUS B101485 481 bp mRNA linear EST 18-JAN-2002
DEFINITION B101485 unpublished oligo-capped cDNA library, C. elegans L1 stage
ACCESSION B101485
VERSION B101485.1 GI:18244155
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Pelodierinae; Caenorhabditis.
1 (bases 1 to 481)
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.
A complementary view of the C. elegans genome
Unpublished
Contact: Tadashi Shin-i
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: shin@genes.nig.ac.jp.
Location/Qualifiers
1..481
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1027a04"
/sex="hermaphrodite"
/issue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

FEATURES

source

1..481
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1027a04"
/sex="hermaphrodite"
/issue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

BASE COUNT 135 a 107 c 104 g 135 t
ORIGIN

Query Match 2.5%; Score 44.2; DB 12; Length 481;
Best Local Similarity 56.6%; Pred. No. 0.27;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

LOCUS B101485 481 bp mRNA linear EST 18-JAN-2002
DEFINITION B101485 unpublished oligo-capped cDNA library, C. elegans L1 stage
ACCESSION B101485
VERSION B101485.1 GI:18244155
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Pelodierinae; Caenorhabditis.
1 (bases 1 to 481)
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.
A complementary view of the C. elegans genome
Unpublished
Contact: Tadashi Shin-i
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: shin@genes.nig.ac.jp.
Location/Qualifiers
1..481
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1027a04"
/sex="hermaphrodite"
/issue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

DB	384	ATTGATGCCATACCGTACAGAGATG	408
RESULT 4			
LOCUS	BJ103567	486 bp	mRNA linear EST 18-JAN-2002
DEFINITION	BJ103567 unpublished oligo-capped cDNA library, C. elegans L1 stage		
ACCESSION	BJ103567		
VERSION	BJ103567.1	GI:18246237	
KEYWORDS	EST.		
SOURCE	Caenorhabditis elegans		
ORGANISM	Caenorhabditis elegans		
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea		
AUTHORS	; Rhabditidae; Pelodierinae; Caenorhabditis.		
TITLE	1 (bases 1 to 486)		
JOURNAL	Kohara, Y., Shin-1, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.		
COMMENT	A complementary view of the C. elegans genome		
FEATURES	Unpublished		
source	Contact: Tadasu Shin-i		
	Center For Genetic Resource Information		
	National Institute of Genetics		
	1111 Yata, Mishima, Shizuoka 411-8540, Japan		
	Tel: 81-559-81-6856		
	Fax: 81-559-81-6855		
	Email: tshin1@genes.nig.ac.jp.		
	Location/Qualifiers		
	1..486		
	/organism="Caenorhabditis elegans"		
	/mol_type="mRNA"		
	/strain="N2"		
	/db_xref="taxon:6239"		
	/clone="YK1051f01"		
	/sex="hermaphrodite"		
	/tissue.type="whole animal"		
	/dev_stage="L1"		
	/clone_lib="unpublished oligo-capped cDNA library, C.		
	elegans L1 stage"		
BASE COUNT	137 a 104 c 114 g 130 t	1 others	
ORIGIN			
Query Match	2.5%;	Score 44.2;	DB 12; Length 486;
Best Local Similarity	56.6%;	Pred. No. 0.28;	
Matches	82; Conservative	0; Mismatches 63; Indels	0; Gaps 0;
Qy	706	AAGCGAAGGCGCGCATTCGGCGCTGCGCACTCGGTGGCAAGAACTG	765
Db	193	AAACCGTTATAGCAATGTCGGCCCGATTTTCGGACTATGCACCGCCGTCGACTG	252
Qy	766	CTTCATGCTGGGGTGAAGCATGTTACATATATATGAAAGCAAGTGAATCGTGTGAGGCAAG	825
Db	253	ATTGAACCTTGAATGACGATTTTGATATCTTACAGAGGTCGACCGGATGGAGGACGA	312
Qy	826	CTTGGTCAATGCTTTTCAAGGACG	850
Db	313	ATTGATGCCATACCGTACAGAGATG	337
RESULT 5			
LOCUS	AU199996	508 bp	mRNA linear EST 17-JUL-2001
DEFINITION	AU199996 unpublished oligo-capped cDNA library, stage L2		
ACCESSION	AU199996		
VERSION	AU199996.1	GI:14827576	
KEYWORDS	EST.		
SOURCE	Caenorhabditis elegans		
ORGANISM	Caenorhabditis elegans		
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea		
AUTHORS	; Rhabditidae; Pelodierinae; Caenorhabditis.		
	1 (bases 1 to 508)		
	Kohara, Y., Shin-1, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.		

[illegible]

```
/clone="yk109b04"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
BASE COUNT      150 a      108 c      117 g      135 t      1 others
ORIGIN

Query Match      2.5%: Score 44.2; DB 12; Length 511;
Best Local Similarity 56.6%; Pred. No. 0.28;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AAGCCGAAGTGGCGTCATTTGGCGTCGATTCGCGACTCGTGTGGCAAGCAACTG 765
DB 209 AAACCGTCTATGCAATGTCGGCGCGGTATTTCCGACTATCGACCGCGCTGCACTG 268
QY 766 CTTCATCTGGCGGTAGACGATGTTACATATATGAAAGTATGCTGTGTGAGCGCAAG 825
DB 269 ATTGAACCTTGGAATTGACGATTTTGATATCTACGAAGTCTCGACCGGATCGAGCGACA 328
QY 826 CTTTGTGCACATGCTTTTCAGGACG 850
DB 329 ATTCATGCCATACCGTACAGGATG 353

RESULT 7
AU203802
LOCUS AU203802 513 bp mRNA linear EST 30-MAY-2003
DEFINITION Caenorhabditis elegans cDNA library, stage L1
ACCESSION AU203802
VERSION AU203802.2 GI:31231885
KEYWORDS EST.
ORGANISM Caenorhabditis elegans
          Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
          1 (bases 1 to 513)
          Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
          and Sugano,S.
          A complementary view of the C.elegans genome
          Unpublished
          On Jul 17, 2001 this sequence version replaced gi:14834726.
          Contact: Yuji Kohara
          Genome Biology Lab.
          National Institute of Genetics
          Yata 1111, Mishima, Shizuoka 411, Japan
          Tel: 81-559-81-6854
          Fax: 81-559-81-6855
          Email: ykohara@lab.nig.ac.jp.
          Location/Qualifiers
            source
              1..513
                /organism="Caenorhabditis elegans"
                /mol_type="mRNA"
                /strain="N2"
                /db_xref="taxon:6239"
                /clone="yk825e11"
                /sex="Hermaphrodite"
                /tissue_type="whole animal"
                /dev_stage="L1"
                /clone_lib="unpublished oligo-capped cDNA library, stage
                L1"
BASE COUNT      141 a      114 c      114 g      143 t      1 others
ORIGIN

Query Match      2.5%: Score 44.2; DB 9; Length 513;
Best Local Similarity 56.6%; Pred. No. 0.28;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AAGCCGAAGTGGCGTCATTTGGCGTCGATTCGCGACTCGTGTGGCAAGCAACTG 765
DB 230 AAACCGTCTATGCAATGTCGGCGCGGTATTTCCGACTATCGACCGGCGCTGCACTG 269
```

```
QY 766 CTTCATCTGGCGGTAGACGATGTTACATATATGAAAGTATGCTGTGTGAGCGCAAG 825
DB 290 ATTGAACCTTGGAATTGACGATTTTGATATCTACGAAGTCTCGACCGGATCGAGCGACA 349
QY 826 CTTTGTGCACATGCTTTTCAGGACG 850
DB 350 ATTCATGCCATACCGTACAGGATG 374

RESULT 8
BU120847
LOCUS BU120847 529 bp mRNA linear EST 23-JAN-2002
DEFINITION BU120847 unpublished oligo-capped cDNA library, C. elegans L1 stage
ACCESSION BU120847
VERSION BU120847.1 GI:18280980
KEYWORDS EST.
ORGANISM Caenorhabditis elegans
          Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
          1 (bases 1 to 529)
          Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
          and Sugano,S.
          A complementary view of the C.elegans genome
          Unpublished
          Contact: Tadasu Shin-I
          Center For Genetic Resource Information
          National Institute of Genetics
          111 Yata, Mishima, Shizuoka 411-8540, Japan
          Tel: 81-559-81-6856
          Fax: 81-559-81-6855
          Email: tshini@genes.nig.ac.jp.
          Location/Qualifiers
            source
              1..529
                /organism="Caenorhabditis elegans"
                /mol_type="mRNA"
                /strain="N2"
                /db_xref="taxon:6239"
                /clone="yk1261a07"
                /sex="hermaphrodite"
                /tissue_type="whole animal"
                /dev_stage="L1"
                /clone_lib="unpublished oligo-capped cDNA library, C.
                elegans L1 stage"
BASE COUNT      154 a      112 c      121 g      142 t

Query Match      2.5%: Score 44.2; DB 12; Length 529;
Best Local Similarity 56.6%; Pred. No. 0.28;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AAGCCGAAGTGGCGTCTCATTTGGCGCTGCGATTCGCGACTCGTGTGGCAAGCAACTG 765
DB 193 AAACCGTCTATGCAATGTCGGCGCGGTATTTCCGACTATCGACCGGCGCTGCACTG 252
QY 766 CTTCATCTGGCGGTAGACGATGTTACATATATGAAAGTATGCTGTGTGAGCGCAAG 825
DB 253 ATTGAACCTTGGAATTGACGATTTTGATATCTACGAAGTCTCGACCGGATCGAGCGACA 312
QY 826 CTTTGTGCACATGCTTTTCAGGACG 850
DB 313 ATTCATGCCATACCGTACAGGATG 337

RESULT 9
AU201678
LOCUS AU201678 540 bp mRNA linear EST 17-JUL-2001
DEFINITION AU201678 unpublished oligo-capped cDNA library, stage L1
ACCESSION AU201678
VERSION AU201678.1 GI:14830776
```

KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS ; Rhabditidae; Pelodierinae; Caenorhabditis.
1 (bases 1 to 540)
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
FEATURES
source
1. 540
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk797d07"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_1lb="unpublished oligo-capped cdna library, stage
L1"
BASE COUNT 157 a 117 c 119 g 147 t
ORIGIN
Query Match 2.5%; Score 44.2; DB 9; Length 540;
Best Local Similarity 56.6%; Pred. No. 0.28;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 706 AACGCCAAGTGGCGGTCATTTGGCGCTGCGATTCGCGCAACGAAGT 765
DB 229 AACCGCTATACCATCGTCGCGCGCGGATTTCCGACATACGACCGCGCTCAGT 288
QY 766 CTTTCATGCTGGGGTAGACGATGTTACATATATGAAGCAAGTATGCTGTGGAGCAG 825
DB 289 ATTGAACCTTGGAATTCGATTTGATATCTACGAAGTCTGCAGCGGATCGAGGCAAG 348
QY 826 CTTTGTCAATGCTTTTCAGGAGC 850
DB 349 ATTCAATGCCATACCGTACAGAGATG 373
RESULT 10
CB401025 560 bp mRNA linear EST 15-MAY-2003
LOCUS OSTP186f11_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
DEFINITION CB401025
ACCESSION CB401025
VERSION CB401025.1 GI:30742752
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS ; Rhabditidae; Pelodierinae; Caenorhabditis.
1 (bases 1 to 560)
Reboul,J., Vaglio,P., Rual,J.F., Lamesch,P., Martinez,M., Armstrong
,C.M., Li,S., Jacotot,L., Berth,N., Janky,R., Moore,T., Hudson
,J.R., Hartley,J.L., Brasch,M.A., Vandenhaute,J., Boulton,S.,
Endress,G.A., Jena,S., Chevot,E., Papasotiriopoulos,V., Tolias,P.P.,
Plescek,J., Snyder,M., Huang,R., Chance,M.R., Lee,H.,
Doucette-Stamm,L., Hill,D.E. and Vidal,M.
TITLE C. elegans ORFome version 1.1: experimental verification of the
genome annotation and resource for proteome-scale protein
expression
JOURNAL Nat. Genet., (2003) In press
COMMENT Contact: Vidal M
Marc Vidal laboratory

Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were
designed on the predicted protein encoding ORF. C. elegans ORFome
cloning project : Contact david_hill@dfci.harvard.edu or
marc.vidal@dfci.harvard.edu
POLYA=No.
FEATURES
source
1. 560
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="Hermaphrodite and male"
/tissue_type="whole animal"
/dev_stage="mixed stage"
/clone_1lb="AD-wrmcDNA"
/note="The AD-wrmcDNA library was generated with poly(A)+
RNA isolated from both hermaphrodite and male N2 worms of
all larval stages, embryos, adults and dauers and the
subsequent generation of cDNAs by poly(A) priming. The
cDNAs were cloned into pC086"
BASE COUNT 175 a 116 c 122 g 147 t
ORIGIN
Query Match 2.5%; Score 44.2; DB 14; Length 560;
Best Local Similarity 56.6%; Pred. No. 0.28;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 706 AACGCCAAGTGGCGGTCATTTGGCGCTGCGATTCGCGCAACGAAGT 765
DB 69 AACCGCTATACCATCGTCGCGCGCGGATTTTCGGAATTCGACCGCGCTCAGT 128
QY 766 CTTTCATGCTGGGGTAGACGATGTTACATATATGAAGCAAGTATGCTGTGGAGCAG 825
DB 129 ATTGAACCTTGGAATTCGATTTGATATCTACGAAGTCTGCAGCGGATCGAGGCAAG 188
QY 826 CTTTGTCAATGCTTTTCAGGAGC 850
DB 189 ATTCAATGCCATACCGTACAGAGATG 213
RESULT 11
BJ121613 584 bp mRNA linear EST 23-JAN-2002
LOCUS BJ121613
DEFINITION BJ121613 unpublished oligo-capped cdna library, C. elegans L1 stage
Caenorhabditis elegans cDNA clone yk1278f01 5', mRNA sequence.
ACCESSION BJ121613
VERSION BJ121613.1 GI:18281750
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS ; Rhabditidae; Pelodierinae; Caenorhabditis.
1 (bases 1 to 584)
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadao Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6855
Fax: 81-559-81-6855
Email: tsuhin@genes.nig.ac.jp.
Location/Qualifiers
FEATURES
source
1. 584
/organism="Caenorhabditis elegans"
/mol_type="mRNA"

[illegible]

Db	210	AAACCGGCTTTAGCAATCGTCGGCGCCGGATTTCCGGACTATCGACCGCGCGTCGACTG	265
Qy	766	CTTCATCCTCGGGGAGACGATGTTACAATATATGAAGCAAGTATCGTGTGGAGGCAAG	825
Db	270	ATTGAACCTTGGAAATTGACGATTTTGATATCTACGAAGGTCTCGACCGGATCGAGGCGACA	329
Qy	826	CTTGTGTCACATGCTTTCAGGACG	850
Db	330	ATTGATGCCATACCGTACAAAGATG	354
RESULT 13			
LOCUS	AU207635	597 bp	mRNA linear EST 17-JUL-2001
DEFINITION	AU207635	unpublished oligo-capped cDNA library, stage L1	
ACCESSION	AU207635		
VERSION	AU207635.1	GI:14841711	
KEYWORDS	EST.		
SOURCE	Caenorhabditis elegans		
ORGANISM	Caenorhabditis elegans		
REFERENCE	Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae		
AUTHORS	1 (bases 1 to 597) Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y. and Sugano,S.		
TITLE	A complementary view of the C.elegans genome		
JOURNAL	Unpublished		
COMMENT	Contact: Yuji Kohara Genome Biology Lab. National Institute of Genetics Yata 111, Mishima, Shizuoka 411, Japan Tel: 81-559-81-6854 Fax: 81-559-81-6855 Email: ykohara@lab.nig.ac.jp Location/Qualifiers		
FEATURES			
source	1..597		
	/organism="Caenorhabditis elegans"		
	/mol_type="mRNA"		
	/strain="N2"		
	/db_xref="taxon:6239"		
	/clone="YK1004b11"		
	/sex="Hermaphrodite"		
	/tissue_type="whole animal"		
	/dev_stage="L1"		
	/clone_idb="unpublished oligo-capped cDNA library, stage L1"		
BASE COUNT	179 a 126 c 131 g 161 t		
ORIGIN			
Query Match	2.5%; Score 44.2; DB 9; Length 597;		
Best Local Similarity	56.6%; Pred. No. 0.29;		
Matches	82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;		
Qy	706	AAGCGGAAGGGGGGCTGATTTGGGGCTGGGCACTTCCGGACTCGTGGTGGCAAAAGCACTG	765
Db	223	AAACCGTCTAAGCAATCGTCGGCGCCGGATTTCCGGACTATGACCGCGCTCGACTG	282
Qy	766	CTTCATCCTCGGGGAGACGATGTTACAATATATGAAGCAAGTATCGTGTGGAGGCAAG	825
Db	283	ATTGAACCTTGGAAATTGACGATTTTGATATCTACGAAGGTCTCGACCGGATCGAGGCGACA	342
Qy	826	CTTGTGTCACATGCTTTCAGGACG	850
Db	343	ATTGATGCCATACCGTACAAAGATG	367
RESULT 14			
LOCUS	BJ101145	603 bp	mRNA linear EST 18-JAN-2002
DEFINITION	BJ101145	unpublished oligo-capped cDNA library, C. elegans L1 stage	
ACCESSION	BJ101145		
	Caenorhabditis elegans cDNA clone YK1022g05 5', mRNA sequence.		

VERSION BJ011145.1 GI:18243815
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans
 Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Pelodierinae; Caenorhabditis.
 1 (bases 1 to 603)
 Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
 and Sugano,S.
 A complementary view of the C.elegans genome
 Unpublished
 Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
 Location/Qualifiers
 1..603
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="YK1022905"
 /sex="hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L1"
 /clone_lib="unpublished oligo-capped cDNA library, C.
 elegans L1 stage"
 BASE COUNT 180 a 130 c 131 g 162 t
 ORIGIN

Query Match 2.5%; Score 44.2; DB 12; Length 603;
 Best Local Similarity 56.6%; Pred. No. 0.29;
 Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAAGTGGCGGTCATTGGCGTGGCATTTCCGACTCGTGTGCAACGAAGT 765
 |||||
 DB 229 AAACCGCTATAGCAATCGTCGGCGCGGTATTTCCGACTATGACCGCGCTCGACTG 288
 |||||
 QY 766 CTTTCATGCTGGGGTAGACGATGTTACATATATGAAGCAAGTATCGTGTGAGGCAAG 825
 |||||
 DB 289 ATTGAACCTTGGAATTGACGATTTTGATATCTACGAAGGTCTGACCGGATCGAGGCAAG 348
 |||||
 QY 826 CTTTGTACATGCTTTTCAGGAGC 850
 |||||
 DB 349 ATTCAATCCATACCGTACAGAGATG 373
 |||||

RESULT 15
 AU204573 608 bp mRNA linear EST 17-JUL-2001
 LOCUS AU204573 unpublished oligo-capped cDNA library, stage L4
 DEFINITION Caenorhabditis elegans cDNA clone yk839b11 5', mRNA sequence.
 ACCESSION AU204573
 AU204573 GI:14836084
 VERSION AU204573.1 GI:14836084
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans
 Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Pelodierinae; Caenorhabditis.
 1 (bases 1 to 608)
 Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
 and Sugano,S.
 A complementary view of the C.elegans genome
 Unpublished
 Contact: Yuji Kohara
 Genome Biology Lab.
 National Institute of Genetics
 Yata 111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855

Email: ykohara@lab.nig.ac.jp.
 Location/Qualifiers
 1..608
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="YK839b11"
 /sex="Hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L4"
 /clone_lib="unpublished oligo-capped cDNA library, stage
 L4"
 BASE COUNT 188 a 129 c 131 g 160 t
 ORIGIN

Query Match 2.5%; Score 44.2; DB 9; Length 608;
 Best Local Similarity 56.6%; Pred. No. 0.29;
 Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAAGTGGCGGTCATTGGCGTGGCATTTCCGACTCGTGTGCAACGAAGT 765
 |||||
 DB 173 AAACCGCTATAGCAATCGTCGGCGCGGTATTTCCGACTATGACCGCGCTCGACTG 232
 |||||
 QY 766 CTTTCATGCTGGGGTAGACGATGTTACATATATGAAGCAAGTATCGTGTGAGGCAAG 825
 |||||
 DB 233 ATTGAACCTTGGAATTGACGATTTTGATATCTACGAAGGTCTGACCGGATCGAGGCAAG 292
 |||||
 QY 826 CTTTGTACATGCTTTTCAGGAGC 850
 |||||
 DB 293 ATTCAATCCATACCGTACAGAGATG 317
 |||||

Search completed: November 24, 2003, 04:01:40
 Job time : 3837.59 secs

THIS PAGE BLANK (USPTO)

Db	5929	GAAGAAGCTTTAGGGGAGGAGGATTACTGAATCTCCACCGAGTGGCGCTGGGTTA	5988
Qy	181	GCTTGCAAAAGGCTGGCCGATGTCGTCTCCCGAGATCTAGCTGGTGGAAAGTGAACA	240
Db	5989	GCTTGCAAAAGGCTGGCCGATGTCGTCTCCCGAGATCTAGCTGGTGGAAAGTGAACA	6048
Qy	241	GTTCTCTCCGCTTATATCTATATATTGGGCAAAATAATCTGGGGGGGAGTATCTTGAAATCGAAA	300
Db	6049	GTTCTCTCCGCTTATATCTATATATTGGGCAAAATAATCTGGGGGGGAGTATCTTGAAATCGAAA	6108
Qy	301	CCTTGGGGCGGGGCAACAGTGAAGTGTCTCGTGGCCATCGACTTGGACACCCATTGTGATG	360
Db	6109	CCTTGGGGGGGGGCAACAGTGAAGTGTCTCGTGGCCATCGACTTGGACACCTTTTGGATG	6168
Qy	361	GATTTCTCCGAGACACACTATCCAGCCCTGTTTTGCTGAGCGGTAAAGATGTGCA	420
Db	6169	GATTTCTCCGAGACACACTATCCAGCCCTGTTTTGCTGAGCGGTAAAGATGTGCA	6228
Qy	421	CCGATTTAGTCTTATGCAATTCGTGGGCCATTTGCATCTCTAGAGCTGCOCGGCTTTCGAAAC	480
Db	6229	CCGATTTAGTCTTATGCAATTCGTGGGCCATTTGCATCTCTAGAGCTGCOCGGCTTTCGAAAC	6288
Qy	481	CTGCCTAATGCCCGCTGTGAAGAGATGGCAGATGAATGCGTTTACCGGGTTTACCATTAAC	540
Db	6289	CTGCCTAATGCCCGCTGTGAAGAGATGGCAGATGAATGCGTTTACCGGGTTTACCATTAAC	6348
Qy	541	CTTGAAGGGGCGGTGCCATTTTGACATGGTAGCTTATGGTCGAAACTTGATGCTGAAGGT	600
Db	6349	CTTGAAGGGGCGGTGCCATTTTGACATGGTAGCTTATGGTCGAAACTTGATGCTGAAGGT	6408
Qy	601	TGGCAGGTTCCCTTCCAAACAATGCACTGTGCTACACACTGACACACGTTTGTGACCAA	660
Db	6409	TGGCAGGTTCCCTTCCAAACAATGCACTGTGCTACACACTGACACACGTTTGTGACCAA	6468
Qy	661	TGTTCCGATAGTGAACGGATGGCGCTTCTTCCGAGAGATGTTCTTAAAGCCGAAAGTGAACG	720
Db	6469	TGTTCCGATAGTGAACGGATGGCGCTTCTTCCGAGAGATGTTCTTAAAGCCGAAAGTGAACG	6528
Qy	721	GTCATTGGCGCTGGCAATTTCCGAACTCGTGGTGGCAAAAGAACTGTTCAATGCTGGGGTA	780
Db	6529	GTCATTGGCGCTGGCAATTTCCGAGAGTGGTGGTGGCAAAAGAACTGTTCAATGCTGGGGTA	6588
Qy	781	GACGATGTTCAAATATATATGAAGCAAGTATGCTGTGGAGGCAAGCTTTGGTGCATGCT	840
Db	6589	GACGATGTTCAAATATATATGAAGCAAGTATGCTGTGGAGGCAAGCTTTGGTGCATGCT	6648
Qy	841	TTCAAGGAGCGCTCCAGTGTGCTGAGCGGAAATGGAGGAGTGGCATTTCCCTCGCTGCA	900
Db	6649	TTCAAGGAGCGCTCTTATGTGTGTGGCCGAAATGGAGGAGTGGCATTTCCCTCGCTGCA	6708
Qy	901	TTCTGCTTGTTTTTCTTCTCGAGCGTTTACGGCGCTGTCTTGATGAGGCGGTTCCCAAT	960
Db	6709	TTCTGCTTGTTTTTCTTCTCGAGCGTTTACGGCGCTGTCTTGATGAGGCGGTTCCCAAT	6768
Qy	961	CCCGGCAACGTCGACACTTATTTGTGTCTTACCAAGCGTCCAAATACATGTGAAAGCCGGG	1020
Db	6769	CCCGGCAACGTCGACACTTATTTGTGTCTTACCAAGCGTCCAAATACATGTGAAAGCCGGG	6828
Qy	1021	CAGCTGCACACGAAAGCTGTTGCATCGCGGTTTAAACAACGGTTGGCGGTGTTTGAAGGAC	1080
Db	6829	CAGCTGCACACGAAAGCTGTTGCATCGCGGTTTAAACAACGGTTGGCGGTGTTTGAAGGAC	6888
Qy	1081	GGTTTTCATGAGCGAGATATTGTGTGGCTTCGCGCTGTGCTATTAATCTCAGGCTTGAAA	1140
Db	6889	GGTTTTCATGAGCGAGATATTGTGTGGCTTCGCGCTGTGCTATTAATCTCAGGCTTGAAA	6948
Qy	1141	TCAGGACACATTAAGTGGGCTCATGACTCTGCGAAATTTGGCTGAACCGTTTGGGAGG	1200
Db	6949	TCAGGAGACATTAAGTGGGCTCATGACTCTGCGAAATTTGGCTGAACCGTTTGGGAGG	7008
Qy	1201	GAGTCTTCTTCTTACAGGATATGAAGATCTTTCTGGGCAACATCTCTCGTGGTGT-AA	1259
Db	7009	GAGTCTTCTTCTTACAGGATATGAAGATCTTTCTGGGCAACATCTCTCGTGGTGTGAA	7068

OY	1260	ACATGAGATTTCCTCATGTATGGGACCTTATCAAGCTAATGGGAATAGGATCTGGCGGG	13139
Dd	7069	ACATGAGATTTCCTCATGTATGGGACCTTATCAAGCTAATGGGAATAGGATCTGGCGGG	71227
OY	1320	GTTCGTGCAGTTTGTGAAAGCGGGTTATTAGATCCTCCGCTTGGTCATCAACGGATA	13799
Dd	7128	GTTCGTGCAGTTTGTGAAAGCGGGTTATTAGATCCTCCGCTTGGTCATCAACGGATA	71879
OY	1380	TGAAGAAAATCAGCGGATGTGTCCCTGAAAGGAATCTCAGAATCTCCAGTCCGATCCGATC	14399
Dd	7188	TGAAGAAAATCAGCGGATGTGTCCCTGAAAGGAATCTCAGAATCTCCAGTCCGATCCGATC	72479
OY	1440	TGAAGTGTTTAAAGGTGTGTCTGTAGACACGCGATATGCAATGCTTCGAAGTCAGGCGAT	14999
Dd	7248	TGAAGTGTTTAAAGGTGTGTCTGTAGACACGCGATATGCAATGCTTCGAAGTCAGGCGAT	73079
OY	1500	TCAGAAAGAAAAGACAAAATTAAGATTAAGGCTTTAAGAGCGGGATATCTGAATCTTTA	15599
Dd	7308	TCAGAAAGAAAAGACAAAATTAAGATTAAGGCTTTAAGAGCGGGATATCTGAATCTTTA	73679
OY	1560	TAAAGTGTTGTCATCTGTGATCTGCAATATCAAACTCAGGCGATTGCTGCATGCGA	16199
Dd	7368	TAAAGTGTTGTCATCTGTGATCTGCAATATCAAACTCAGGCGATTGCTGCATGCGA	74279
OY	1620	TACCAATATTTTTTAAAGGACACAGTGAACCAAGCGGTGATTAACAGCCATATGACAGATC	16779
Dd	7428	TACCAATATTTTTTAAAGGACACAGTGAACCAAGCGGTGATTAACAGCCATATGACAGATC	74879
OY	1680	GTCAAAACTCTTCTCTGATGACTGAACGAAAATTTGTGTTAGACCATATCTCCCGTCTTG	17399
Dd	7488	GTCAAAACTCTTCTCTGATGACTGAACGAAAATTTGTGTTAGACCATATCTCCCGTCTTG	75479
OY	1740	TGTTCCTCATGAGCGGATCCGCAAAAGCATGTATTCCTGTGACTATGATGTGACAGATCC	17999
Dd	7548	TGTTCCTCATGAGCGGATCCGCAAAAGCATGTATTCCTGTGACTATGATGTGACAGATCC	76079
OY	1800	GA 1801	
Dd	7608	GA 7609	

RESULT 2
US-09-073-587-2
Sequence 2, Application US/09073587
Patent No. 6268552
GENERAL INFORMATION:
APPLICANT: Li, Yi
TITLE OF INVENTION: Transgenic Seedless Plants
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5170 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,587
FILING DATE: 06-MAY-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/045,725
FILING DATE: 06-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Feather, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 4-97

TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2211 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 4..2205
US-09-073-587-2

Query Match 94.2%; Score 1697.4; DB 3; Length 2211;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1726; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY 64 ACAATGTCATTAAGCGGATGATTTGACCGCGGTTTCCGATGCTTCTTGAACGA 123
DB 1 ACAATGTCATTAAGCGGATGATTTGACCGCGGTTTCCGATGCTTCTTGAACGA 60
QY 124 GAAGCTCTAGGGGAGAGGATTAACCAAGCTCCAGTGCAGCGTGGGTTAGCT 183
DB 61 GAAGCTCTAGGGGAGAGGATTAACCAAGCTCCAGTGCAGCGTGGGTTAGCT 120
QY 184 TGCAAAAGGCTGCGATGCTGCTCCGAGATCTCAGTGTGAGAAAGTAGCAGT 243
DB 121 TGCAAAAGGCTGCGATGCTGCTCCGAGATCTCAGTGTGAGAAAGTAGCAGT 180
QY 244 CTCTCCGCTTATATCTATTTGGCAAGAAATTTGGGGGAGATCTTGAATCGAACT 303
DB 181 CTCTCCGCTTATATCTATTTGGCAAGAAATTTGGGGGAGATCTTGAATCGAACT 240
QY 304 TGGGCGGGGCAACAGAGTGTCTGTTGCCATGCACTTGACACCAATTTGCATGAT 363
DB 241 TGGGCGGGGCAACAGAGTGTCTGTTGCCATGCACTTGACACCAATTTGCATGAT 300
QY 364 TTCTCCGACACACATTAATCCAGCCCTGTTTGTGAGCGGTAAAGATGACACG 423
DB 301 TTCTCCGACACACATTAATCCAGCCCTGTTTGTGAGCGGTAAAGATGACACG 360
QY 424 ATTGATCTTAGTCACTTTGCGGCATTTCAATCTTAAGACTGCGGCTTTCGAACTCG 483
DB 361 ATTGATCTTAGTCACTTTGCGGCATTTCAATCTTAAGACTGCGGCTTTCGAACTCG 420
QY 484 CCAATGCGGCTGACGAGATGCGAGATGAAATGCGCTTACCGGTTTACCATACCTT 543
DB 421 CCAATGCGGCTGACGAGATGCGAGATGAAATGCGCTTACCGGTTTACCATACCTT 480
QY 544 GAAGGGGCGTGCATTTGACATGTTAGTGTGAGAAACCTGATGCTGAAGGTTTCG 603
DB 481 GAAGGGGCGTGCATTTGACATGTTAGTGTGAGAAACCTGATGCTGAAGGTTTCG 540
QY 604 GCAAGTTCTTTCCACCAATGCACTTCTCAAGCTGACAGCCGTTTTTGAACCAATGT 663
DB 541 GCAAGTTCTTTCCACCAATGCACTTCTCAAGCTGACAGCCGTTTTTGAACCAATGT 600
QY 664 TCCGATGTCGAGGAGATCGGCTTCTTCCGAGAGATTTCTTAAAGCGGAAGTGGCGTC 723
DB 601 TCCGATGTCGAGGAGATCGGCTTCTTCCGAGAGATTTCTTAAAGCGGAAGTGGCGTC 660
QY 724 ATTGCGCTGCAATTTCCGAGCTGTTGGGCAACCAAGCTGCTTCAATGCTGGGTTAGC 783
DB 661 ATTGCGCTGCAATTTCCGAGCTGTTGGGCAACCAAGCTGCTTCAATGCTGGGTTAGC 720
QY 784 GATGTTTCAATATATGAGCAAGTATGCTGTTGAGGCAAGCTTTGGTCAATGCTTTC 843
DB 721 GATGTTTCAATATATGAGCAAGTATGCTGTTGAGGCAAGCTTTGGTCAATGCTTTC 780

QY 844 AGGAGCGCTCTAGTGTGCGGCCGAATGAGGGGCGATGCGATTTCTCCTGTCGATTC 903
DB 781 AGGAGCGCTCTAGTGTGCGGCCGAATGAGGGGCGATGCGATTTCTCCTGTCGATTC 840
QY 904 TGCCTGTTTTCTTCTCGAGCGTTACCGCTGTCTTTCGATGAGCGGTTCCCAATCCC 963
DB 841 TGCCTGTTTTCTTCTCGAGCGTTACCGCTGTCTTTCGATGAGCGGTTCCCAATCCC 900
QY 964 GGCACAGTCACACTTACTTGTCTACCAAGGCGTCCAAATCATGTGGAAGCGGGCAG 1023
DB 901 GGCACAGTCACACTTACTTGTCTACCAAGGCGTCCAAATCATGTGGAAGCGGGCAG 960
QY 1024 CTCACACCGAGCTGTTCCATCGCTTTTACCAAGGTTGCGGCTGTTCTTGAAGACGGT 1083
DB 961 CTCACACCGAGCTGTTCCATCGCTTTTACCAAGGTTGCGGCTGTTCTTGAAGACGGT 1020
QY 1084 TTTCAATGAGGAGATATGTTGTTGGCTGCTGTGCTATTAATCTAGGCTTGAATCA 1143
DB 1021 TTTCAATGAGGAGATATGTTGTTGGCTGCTGTGCTATTAATCTAGGCTTGAATCA 1080
QY 1144 GGCACATTAAGTGGGCTCATGCTCTGGCAAAATTTGGCTGAACCGTTTCGGAGGAG 1203
DB 1081 GGCACATTAAGTGGGCTCATGCTCTGGCAAAATTTGGCTGAACCGTTTCGGAGGAG 1140
QY 1204 TCTCTCTCTCAGGATAGAGAGATCTTCTGGGACACATCTCTGCTGGTGGT-AAACA 1262
DB 1141 TCTCTCTCTCAGGATAGAGAGATCTTCTGGGACACATCTCTGCTGGTGGTAAACA 1200
QY 1263 TGGAGTTTCTCATGATTTGGGACCTATTCAAGCTAATGGAATAGAGATGCGGGGTT 1322
DB 1201 TGGAGTTTCTCATGATTTGGGACCTATTCAAGCTAATGGAATAGAGATGCGG 1259
QY 1323 TGGTCAAGTTTGAAGCGGTTTATTTAGATCTCTCGCTTGGTCAACGAGATGCA 1382
DB 1260 TGGTCAAGTTTGAAGCGGTTTATTTAGATCTCTCGCTTGGTCAACGAGATGCA 1319
QY 1383 AGAAATTCAGCGGATGTCCTTGAAGAAATCTGAGAACTTCCAGCTCGGATGCGATCTGA 1442
DB 1320 AGAAATTCAGCGGATGTCCTTGAAGAAATCTGAGAACTTCCAGCTCGGATGCGATCTGA 1379
QY 1443 AGTGTATTAACGATGTCGTGAGCCAGCGCATATGCTTCAAGTCAAGGCGATTC 1502
DB 1380 AGTGTATTAACGATGTCGTGAGCCAGCGCATATGCTTCAAGTCAAGGCGATTC 1439
QY 1503 GAAGGAAAGACAAAATTAAGATTAAGCTTAAAGACCGGATATCTGAACCTTATGATTA 1562
DB 1440 GAAGGAAAGACAAAATTAAGATTAAGCTTAAAGACCGGATATCTGAACCTTATGATTA 1499
QY 1563 GGTGTGTGTCACATCGGACCTGCAATATCCAACTCAGGCAATGCTGACATGCGATAC 1622
DB 1500 GGTGTGTGTCACATCGGACCTGCAATATCCAACTCAGGCAATGCTGACATGCGATAC 1559
QY 1623 CAATATTTTTCAGGACACAGTGAACCAAGCGGTTGTTAAACAGCATATGACAGATCTC 1682
DB 1560 CAATATTTTTCAGGACACAGTGAACCAAGCGGTTGTTAAACAGCATATGACAGATCTC 1619
QY 1683 AAAACTCTTCTGATGATGTAACGAAATTTCTGTTTGAACCATATCTCTCCGTTGTGT 1742
DB 1620 AAAACTCTTCTGATGATGTAACGAAATTTCTGTTTGAACCATATCTCTCCGTTGTGT 1679
QY 1743 CCTCATGAGCGGATGCAAAAGCAAGTATTTGCCGTCATATGATTCAGAGATCCGA 1801
DB 1680 CCTCATGAGCGGATGCAAAAGCAAGTATTTGCCGTCATATGATTCAGAGATCCGA 1738

RESULT 3
US-08-008-216-8
; Sequence 8, Application US/08008216
; Patent No. 536887
; GENERAL INFORMATION:
; APPLICANT: Slightom, Jerry L.
; APPLICANT: Tepfer, David A.
; TITLE OF INVENTION: R1 T-DNA Promoters

NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: GRAY, CARY, AMES & FRYE
STREET: 401 B Street, Suite 1700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-4297
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/008,216
FILING DATE: 25-JAN-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/725,368
FILING DATE: 22-APR-1985
ATTORNEY/AGENT INFORMATION:
NAME: Barnhorst, Marlene W.
REGISTRATION NUMBER: 36,740
REFERENCE/DOCKET NUMBER: P1020US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 699-2700
TELEFAX: (619) 236-1048
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2279 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Agrobacterium rhizogenes
STRAIN: Strain A4
IMMEDIATE SOURCE:
LIBRARY: Convolulus arvensis plant cells
CLONE: Clone 7
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..2279
OTHER INFORMATION: /label= ORF8
OTHER INFORMATION: /note= "Sequence ORFs corresponds to bases 6609
; OTHER INFORMATION: through 8868 of Seq. ID No. 536887 19."
; US-08-008-216-8
Query Match 6.8%; Score 121.6; DB 1; Length 2279;
Best Local Similarity 45.7%; Pred. No. 1,5e-32;
Matches 540; Conservative 0; Mismatches 634; Indels 8; Gaps 3;
QY 623 TCAGCTTGCTCAGACTGACGACCGTTTTCGACCAATGTTCCGATAGTGAAGATCG 682
DB 665 TTGACCTGCGGTAAGACTATCAGGTTTCTCCAGAGTGCAGTCCCATGATGATGTGG 724
QY 683 GCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGGTCATTGGCGCTGGCATTTCCG 742
DB 725 GGCATTATCCGGAAGGGGCAACACTACTAGTGTGCATCTGTAGAGAGGGGCGCTGTCTG 784
QY 743 GACTCGGTGTGCAAGCAAGCTGCTCATGCTGGGGTAGAGCATGTACAAATATATGAG 802
DB 785 GCCTTGTGCTGCACAGAACTACTTGGCGCTGCGTCAAGGAATACCTCTTTTGATA 844
QY 803 CAAGTATCGTGT--TGAGGCAAGCTTTGTCACATGCTTTTCAGGGACGCTCTTAGT 859
DB 845 CCGTGTATGAGATCCGTAAGTTTGGGGCATGCGCATGCCAAGCGGAGCGCTCACAGG 904
QY 860 TCGTGGCCGAAATGGGGCGCATGCGATTTCTCTGCTGCATTTCTGTTTCTTCTCC 919
DB 905 CTTTGAAGTGTGTCGGGTGTCTTCTTCCGCAACAACTTTGCTCTCATATCATC 964
QY 920 TCGAGCGTTACGGGCTGTCTTCATGAGGCGTTCCCAATCCCGGACAGTGAACATT 979

DB 965 TGATATAGTTTAAATTCGCTCCAGCTTCTTCTTCTGTGCGGCAAGCACACAG 1024
QY 980 ACTTGTCTTACCAAGCGCTTCAATACATGTGAAAGCCGGGACGTGCCACCAAGCTGT 1039
DB 1025 CACTATATTTCCCGCAGAAACGCTACGATGCGACGGGGCAAGCTCCGCGGGGATAT 1084
QY 1040 TCCATCGCGCTTTCACACGGTGGCGGCTGCTTTCGAAAGACCGCTTTCATGACGATA 1099
DB 1085 TTACGGGGGACATGTCGATGGAAGACACTACTTACCAAGGGGTGAAAGGAAATGCA 1144
QY 1100 TTGTGTGGCTTCCGCTGCTGCTATTAATCAGCGCTTGAATCAGGACATTAAGTGGG 1159
DB 1145 GAGACTGATGCTGCTCGATGATATCTTTTCAATGTTAAAGCGCTGCGATGAAG 1204
QY 1160 CTGATACCTCTGGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTTTCAGGGA 1219
DB 1205 CCTCAAGACACGCGAGCTTGGCTCCGAGAGTTCGAAATTCACCTTTCATGCGGTTT 1264
QY 1220 TAGAGAGATCTTCTTGGGCACACATCCCTCGTGTGT--AAACATGAGTTTCTTC 1275
DB 1265 TGATCAGATCTTCAAGCTGTGTAATTCAGTCTGTGCAAGGACATGCAAAACACCC 1324
QY 1276 ATGATGGGACCTATTCAAGCTAATGGAATAGATCTGGCGGGTGTGTCAGTTT 1335
DB 1325 ATGATTTGAGGCTTTGGGATATCTGAGGTGGGATACGGCGAGTTTGTCTC-TATTAC 1383
QY 1336 GAAAGCGGTTTATTAAGATCTCCGCTGTGTGTCATCAACGATATGAAGAAATCAGCGG 1395
DB 1384 AACGTGTTGTTTCAACGATCCGAGCTGATTAATCAATGCTACAGAGGAGACAGCAT 1443
QY 1396 ATGCGCCCTGAAGGAATTCAGAACTTCACAGCTGCGATGCGATCTGAAGGCTTAACGT 1455
DB 1444 CTTCATTGTTGGGGTTCACTTTTGCAGGCTCTATGCGCAATTGAATATTCAGAAA 1503
QY 1456 GTGCTGTGAGCCAGCGCATATGCAATGTTCAAGTCAAGCGGCGATTCAGAAAGAAAGACA 1515
DB 1504 AGCCATGCCAAGAACAGACTCTGTTTGAATCCGTCGCGTGAATATACCAAGAGGGCGG 1563
QY 1516 AAAATTAAGATTAAGCTTAAAGCGGGAATTCGAATCTTATGATGAAGTGTGTACACA 1575
DB 1564 AGATTGAAGTATGCTTGAACACGCTCATTCGCGTGTGTTTGAACAGGCTCATATGAGC 1623
QY 1576 TCTGACTGGCAATATCCAACTCAGGCAATGCTGCAATGCGATACCAATATTTTCAG 1635
DB 1624 GCGAGTCTGAGGCGCTTACAGTTGATTAACAGCTGCGGGGATGAGACTTCTTCAGC 1683
QY 1636 GCACAGTGAACCAAGCGGTTGATTAACAGCATATGACAGATCGTCAAAACTCTTCTG 1695
DB 1684 TACAATATGAAACCGCGCTGCGAAACTGTGCGGCTGTCAATTCAGACTCTTCATG 1743
QY 1696 ATGACTGAACGAAATTTGTGTTAGACATATTCCTCCGCTTGTGTCTCATGACGGG 1755
DB 1744 GTACCAAGCAAAAGTTTGGGTTAACTCCGGCATCCAGCACTGATATGAGACCGATGG 1803
QY 1756 ATGCAAAAGCAAGTATGCTGACTGATGAGTGCAGGAT 1797
DB 1804 CTGTGCTGAGCTGTGTTGATGATGATGATGCGCAGCT 1845
RESULT 4
US-08-459-569-8
; Sequence 8, Application US/08459569
; Patent No. 5543501
; GENERAL INFORMATION:
; APPLICANT: Slightom, Jerry L.
; APPLICANT: Tepler, David A.
; TITLE OF INVENTION: R1 T-DNA Promoters
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GRAY, CARY, AMES & FRYE
; STREET: 401 B Street, Suite 1700
; CITY: San Diego

STATE: California
COUNTRY: USA
ZIP: 92101-4297
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,569
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,216
FILING DATE: 25-JAN-1993
APPLICATION NUMBER: US 06/725,368
FILING DATE: 22-APR-1985
ATTORNEY/AGENT INFORMATION:
NAME: Barnhorst, Marnie W.
REGISTRATION NUMBER: 36,740
REFERENCE/DOCKET NUMBER: P1020U51
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 699-2700
TELEFAX: (619) 236-1048
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2279 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
ORIGINAL SOURCE:
ORGANISM: Agrobacterium rhizogenes
STRAIN: Strain A4
IMMEDIATE SOURCE:
LIBRARY: Convolvulus arvensis plant cells
CLONE: Clone 7
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..2279
OTHER INFORMATION: /label= ORF8
OTHER INFORMATION: /note="Sequence ORF8 corresponds to bases 6609
Patent No. 5543501
OTHER INFORMATION: through 8888 of Seq. ID No. 5543501 19. "
US-08-459-569-8

Query Match 6.8%; Score 121.6; DB 1; Length 2279;
Best Local Similarity 45.7%; Pred. No. 1.5e-32;
Matches 540; Conservative 0; Mismatches 634; Indels 8; Gaps 3;

QY 623 TCGACTTGTCTACGACTGACAGCCGTTTGTGACCAATGTTCCGATGAGCGAGATCG 682
DB 665 TTGACCGCCGGTACGACTATCAGGTTTCTCCAGAGTCCGATGCCATGATGATG 724
QY 683 GCTTCTTCCGAGGAGTTCCTTAAGCCGAAAGGGGGGTCATTGGCGCTGGCATTTCCG 742
DB 725 GGCATTATCCGAGAGGCGCACCACTACTAGTGTGTCATGTGAGGAGGGGCTGTCTG 784
QY 743 GACTCGTGTGCGAAAGCACTGCTTCATCTGGGAGTAGACGATGTACAAATATAGAA 802
DB 785 GCCTGTGTGTCGACAGAACTACTTGGCGCTGGCGTCAAGAAATCACTTTTGGATA 844
QY 803 CAAGTATCTGT--TGAAGGCAAGCTTTGTTCATGCTTTTCAAGGAGCTTCTAGTG 859
DB 845 CCGTGTATGAGATCCGATGTTTGGGCGATCGCCGATGCCAAACGGGAGCTCACAGG 904
QY 860 TCGTGGCGGAAATGGGGGCGATGATTTCTCTCTGTGATCTGCTTTTCTTTC 919
DB 905 CCTTGAGCTGTTGGTGTATGCTTCTTCCGCAACAACCTTGTCTTCAATATC 964
QY 920 TCGAGCGTTACGGCTGTCTTTCATGAGGCGCTTCCCAATCCCGGACAGTGACACTT 979
DB 965 TGGATAGTTTAGAATCCGTCACGCTTGTTTCTTGTGGCGGCAAGACACACAG 1024

QY 980 ACTTGGCTTACCAAGGCGTCCATACATGTGAAAGCCGGGACGTCGCCAGAGCTGT 1039
DB 1025 CACTATATTTCCGCCAAGAAACGCTACGCATGGCAGCGGGCAAGCTCCGCGGGATAT 1084
QY 1040 TCCATCGCGTTTACCAACGCGTGGCGCTTCTTGAAGACGCTTTTCAATGACGAGATA 1099
DB 1085 TTCAGGGGGTACATGTCGATGGAAGACACTACTACCAAGGCTTGAACGGAATGGCA 1144
QY 1100 TTGTGTGGCTTCCGCTGTGCTTATTAAGCCCTTGAATAGACACATTAGTGGG 1159
DB 1145 GGAGCTGATGGCTCCGATGATATCTTTTCAATGTTGAAGAGCCGTCGTGATGAAG 1204
QY 1160 CTGATACCTCGGCAAAATTTGGCTGAACCGTTTGGGAGGGAGTCTTCTTCAGGGA 1219
DB 1205 CCTCAGAGACGGCAGCTTTGGCTCCGAGAGTTCCGAAATTCATTTTCATCCGCTTT 1264
QY 1220 TAGAGAGATCTTCTTGGGACACATCTCTGTGTGT----AAACATGAGTTTCTTC 1275
DB 1265 TGGTCAAGATCTTCACTGTGTGTAATTCAGTCTGTGTGGCAAGGATGGCAAAACCCC 1324
QY 1276 ATGATTGGACCTATTCAAGCTAATGGAATGAGATCTGGCGGGTTTGGTCCAGTTT 1335
DB 1325 ATGATTTGAGGCTTCCGGATCTAGAGGTGGGATACGGCCGAGTTTGTCTC-TATTAC 1383
QY 1336 GAAAGGGGTTTATTAAGATCTCTCCGCTGTGTCAATCAAGGATATGAAGAAATCAGCG 1395
DB 1384 AAGGTGTGTTTCAACGATCTCGAGCTGATTAATGCTACAGAGAGGACGAGCAT 1443
QY 1396 ATGTGCTTGAAGGATCTCAGAACTTCCACGTCGATCGCATCTGAAGTGTTAACGGT 1455
DB 1444 CTTCTTATGTGTGGGTTCACTTTTGCAGGCTGTGATGCGATTGAATATTTCCGAAA 1503
QY 1456 GTGTCTGTGAGCCAGCGCATATGCAATGTTCAAGTCAGGCGCATTCAGAGAAAGACA 1515
DB 1504 AGCCATGCGAAAGACAGACTGTGTTTGTATCCGCTGTGAATACCAAGAGGCGCGG 1563
QY 1516 AAAATTAAGATAAGCTTAAAGCGGGATATCTGAACCTTTATGATAAGTGTGTGACCA 1575
DB 1564 AGATTGAAGTATGCTTGAACACGCTCATTTGCGGTGTTTGAACGAGGTATCATTTG 1623
QY 1576 TCTGACCTGCAAAATATCAACTCAGCATGTCGATGCGATGCAATATTTTTCAG 1635
DB 1624 GGCAGTCTGAGGCGCTCACTTATTAACAGACTGCGGGATAGACTTCTTTCAGC 1683
QY 1636 GCACCAAGTACCAACGCGGTGATTAACAGCAATATGACAGATCTGCAAACTTTCTG 1695
DB 1684 TACAATATGAAACCGCGCGTGAAGAACTGCTGCGGTGTCAATTCAAGACTCTTCATG 1743
QY 1696 ATGACTGAACGAAATTTCTGTTAAGCAATATCTTCCGCTTGTGTCTCATGACGGG 1755
DB 1744 GTCAACGAAACAAAGTTTGGGTAACTCCGATCCGACAGTGAATGAGACCGATGGG 1803
QY 1756 ATCGCAAAAGCAATGATGCTGATGCTGATGATGATGATGATGATGATGATGAT 1797
DB 1804 CTGTGCTGAGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1845

RESULT 5
US-08-458-831-8
Sequence 8, Application US/08458831
Patent No. 582486
GENERAL INFORMATION:
APPLICANT: Slightom, Jerry L.
APPLICANT: Teifer, David A.
TITLE OF INVENTION: R1 T-DNA Promoters
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: GRAY, CARY, AMES & FRYE
STREET: 401 B Street, Suite 1700
CITY: San Diego
STATE: California
COUNTRY: USA

ZIP: 92101-4297
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,831
FILING DATE: 02-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,216
FILING DATE: 25-JAN-1993
APPLICATION NUMBER: US 06/725,368
FILING DATE: 22-APR-1985
ATTORNEY/AGENT INFORMATION:
NAME: Barnhorst, Marlene W.
REGISTRATION NUMBER: 36,740
REFERENCE/DOCKET NUMBER: P1020US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 699-2700
TELEFAX: (619) 236-1048
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2279 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Agrobacterium rhizogenes
STRAIN: Strain A4
IMMEDIATE SOURCE:
LIBRARY: Convolutus arvensis plant cells
CLONE: Clone 7
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..2279
OTHER INFORMATION: /label= ORF8
Patent No. 5824866
OTHER INFORMATION: through 8888 of Seq. ID No. 5824866 19. "
US-08-458-831-8
Query Match 6.8%; Score 121.6; DB 1; Length 2279;
Best Local Similarity 45.7%; Pred. No. 1.5e-32;
Matches 540; Conservative 0; Mismatches 634; Indels 8; Gaps 3;
QY 623 TCAGCTGCTCTAGCTGAGACCGTTTTCACATGTTCCGATAGTGAAGCATCG 682
DB 665 TTGACCTGGGTAGACTATCAGTTTCTCCAGAGTGGAGTCCCATGATGCTGG 724
QY 683 GCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCGTCAATTGGCGCTGCA 742
DB 725 GGCATTATCCGGAAGGCGACACTACCTAGTGTGTCATGTGAGGAGGCGGCTGTG 784
QY 743 GACTCGGTGGGCAAGCACTGCTCATGCTGGGGTGAAGCATGTAAATATAGAG 802
DB 785 GCTTGTGCTGCACAGAACTTGTGCGTGGCGTCAAGAAATATCTTTTCGATA 844
QY 803 CAAGTATCGTGT--TGAGAGCAAGCTTTGTGCATGCTTTCAGGAGCGTCTAGTG 859
DB 845 CCGTGTATGAGATCCGATAGTTTGGGCGATCGCGAATGCCAAGCGGACGCTACAG 904
QY 860 TCGTGCGCAATGGGCGCATGCGATTTCTCTGCTGCATCTGCTGTTTCTTCC 919
DB 905 CTTGAGGTGCTGCTGCTATGCTTCTCCGCAACCACTTGTGCTGCTATCATATC 964
QY 920 TCGAGCGTTACGGCGCTCTTTCATGAGCGCTTCCCAATCCGCGACAGTGCACCT 979
DB 965 TGAATTAAGTTTGAATTCCTGCGAGCTTCTGTTTCTTGGCGGAGACGACACAG 1024
QY 980 ACTTGCTACCAAGCGGTCAATATACATGTAAGCAAGCGGAGAGCTGCCACGAACTGT 1039

DB 1025 CACTATATTTCCGCCAGAAACGCTACGATGACCGCGGGCAAGCTCCGCCGGGATAT 1084
QY 1040 TCCATGCGCTTTACACAGGTTGGCGTCTTCTGTAAGACGGTTTTCATGACCGATA 1099
DB 1085 TTGAGGGGTACATGTCGATGGAAGACACTACTTCAACAAAGGTTGAAAGCAATGCA 1144
QY 1100 TTGTGTGCTTGGCTGCTGCTATTTACTCAGGCTTGAATATGAGACATTAAGTGGG 1159
DB 1145 GAGACTGATGCTCGATGATGATCTTTTCAATGTTAAAGAGCGTCCGATGAGAA 1204
QY 1160 CTCATGATCTCTGGCAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTTCAGGGA 1219
DB 1205 CCTCAAGACGCGCAGCTTGTGCTCCGAGATTCGAAATTCATTTCCAGCCGTTT 1264
QY 1220 TAGAGAGATCTTTTGGGCAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1275
DB 1265 TGGTCAAGATCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1324
QY 1276 ATGATTGGGACCTATTCAAGCTATGGAATAGATCTGCGGGGTTTGGTCCAGTTT 1335
DB 1325 ATGATTTCAAGCTTTCGGGATCTGAGTTGAGATACGCGGAGTTTGTCTC-TATTAC 1383
QY 1336 GAAAGCGGTTTATTAAGATCTCCGCTTGTCTATCAAGATTAAGAAATCAAGCGG 1395
DB 1384 AACGTGTTGTTTCAACGATCTGCACTGATTTATCAATGCTACGAGGAGACAGCAT 1443
QY 1396 ATGCGCCCTGAAGGAATCTCAAGATCTTCAAGCTGCAATGCAATGAGTTTACGCT 1455
DB 1444 CTTCATTATGTTGGGTTCACTTTTGCAGGCTCTATGCGCAATTAATATTCAGAAA 1503
QY 1456 GTGCTGTGAGCGCAGCGCATATGCTATGTTCAAGTCAAGGCGCATTAAGAAAGACA 1515
DB 1504 AGCCATGCGAAAGCAAGACTCTTTTGAATCCGCTGCTGATTAACCAAGAGGCGGG 1563
QY 1516 AAAATTAAGATAAGCTTAAAGCGGATATCTGAATCTTATGATTAAGTGTGTCACA 1575
DB 1564 AGATTGAAGTATGCTTGAACACAGGCTCATTCGCTGTTTGAACAGGCTCATCATGCGC 1623
QY 1576 TCTGAGCTGCAAAATATCAACTCAGGCAATGCTGCTGCAATGCAATATTTTCAG 1635
DB 1624 GCGAGCTGTAGGCGCCCTACAGTTGATTAACAGACTGCGGGATGATGATCTCTTACG 1683
QY 1636 GCACAGTGAACCAAGCGGTTGATTAACAGCATATGACAGATCTCAAAACTCTTCTG 1695
DB 1684 TAAATATGAACCCCGCTCGGAAATCTGCTCGCGCTGTCAATTCAGACTCTTCATG 1743
QY 1696 ATGACTGAACGAAATTTGTGTTAGACATATCTCCGCTTGTGCTCATGACGGG 1755
DB 1744 GTACGAGCAAGAAATTTGGGTTAACTCCGCGATCCAGCATGATGATGACCGATGGG 1803
QY 1756 ATGCAAAAGCACTGATTTGCTGCACTATGATCCAGAT 1797
DB 1804 CTGTCCGTGAGCTGTGTGATTAACATGCAATCCAGCT 1845
RESULT 6
US-08-008-216-19
Sequence 19, Application US/08008216
Patent No. 5366887
GENERAL INFORMATION:
APPLICANT: Slightom, Jerry L.
APPLICANT: Tepfer, David A.
TITLE OF INVENTION: R1-T-DNA Promoters
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: GRAY, CARY, AMES & FRYE
STREET: 401 B Street, Suite 1700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-4297
COMPUTER READABLE FORM:

	NAME/KEY:	misc feature	
?	LOCATION:	complement (10509..11282)	
?	OTHER INFORMATION:	/label= ORF11SUBSEQUENC	
?	FEATURE:		
?	NAME/KEY:	misc feature	
?	LOCATION:	12466..13002	
?	OTHER INFORMATION:	/label= ORF12SUBSEQUENC	
?	FEATURE:		
?	NAME/KEY:	misc feature	
?	LOCATION:	13723..14319	
?	OTHER INFORMATION:	/label= ORF13SUBSEQUENC	
?	FEATURE:		
?	NAME/KEY:	misc feature	
?	LOCATION:	15659..16210	
?	OTHER INFORMATION:	/label= ORF14SUBSEQUENC	
?	FEATURE:		
?	NAME/KEY:	misc feature	
?	LOCATION:	complement (16517..17545)	
?	OTHER INFORMATION:	/label= ORF15SUBSEQUENC	
?	FEATURE:		
?	NAME/KEY:	misc feature	
?	LOCATION:	complement (17737..18189)	
?	OTHER INFORMATION:	/label= ORF16SUBSEQUENC	
?	FEATURE:		
?	NAME/KEY:	misc feature	
?	LOCATION:	complement (18177..18743)	
?	OTHER INFORMATION:	/label= ORF17SUBSEQUENC	
?	FEATURE:		
?	NAME/KEY:	misc feature	
?	LOCATION:	complement (19031..19380)	
?	OTHER INFORMATION:	/label= ORF18SUBSEQUENC	
?	US-08-008-216-19		
Query Match	6.8%;	Score 121.6;	DB 1; Length 21126;
Best Local Similarity	45.7%;	Pred. No. 8.4e-32;	
Matches 540;	Conservative 0;	Mismatches 634;	Indels 8; Gaps 3;
QY	623	TCGACCTGCTCTACGACCTGACGACCGTTTTTTTACCATGTTCCGATAGTGAACGATCG	682
DB	7273	TTGACCGCGGGTACGATACAGGTTTCTCCAGAGTGGATGCCATGATGTTGG	7332
QY	683	GCTTCTTTCCGAGAGATGTTCTTAAAGCCGAAAGTGGCGGTATTTGGCGTGGCATTTCCG	742
DB	7333	GGCATTTACCGGAAGCGCACCACTACCTAGTGTTCATCTGTAGAGCGGGCTGTCTG	7392
QY	743	GACCTCGGTGGCAACGACCTGTTTACCTGTTGGGGTGAACGATTAACAATATATGAG	802
DB	7393	GCCCTGTTGTGTCACAGAACTACTTGGCGCTGGCTGCAAGAAATCACTTTTTCATTA	7452
QY	803	CAAGTGAATCTGT---TGAGGCAAGCTTTGGTACATGCTTTCAAGGAGCGCTCTAGTG	859
DB	7453	CCGTTATGATGATCCGATGTTTGGGGCATCGCGCAATCCAAACGGGACGCTACCAAG	7512
QY	860	TCGTGGCCGAATGGGGCGCATGCGATTTTCCTCTGCTGCATTTGCTTGTGTTTCTTCC	919
DB	7513	CTTTGACGTGCTTCGGGTGTCATGCTTTTCCGCAACCAACTTTGCTGTCACTATAC	7572
QY	920	TCGAGCGTTACGGCGCTGTCTTCGATGAGCGCGTTCCCAATCCCGGACAGTGCACACTT	979
DB	7573	TGGAATAGTTTAAATTTCCGTCACAGCCCTTGTTTCTTGTGCGGGAAGCACACAG	7632
QY	980	ACTTGGCTACCAAGGCGTCCAAATACATGTGGAAGCGGGGACGCTGCAACGAAGCTGT	1039
DB	7633	CACATATATTTCCCGCACAAACGCTACGATGCAACGCGGGGACAGCTCCGCGGGGATAT	7692
QY	1040	TCCATGCGCTTTACCAACGTTGCGTGGCTGCTTTTGAAGACGCTTTTCATGAGCAGATA	1099
DB	7693	TTTACGGGGGATACATGTTCGATGGAAGACATACTTACCAAGGGTGTGAACGGAATGGA	7752
QY	1100	TTGTGTTGGCTTGGCGCTGTGCTATTAACAAGCGCTTGAATCAGGACACATTAAGTGGG	1159
DB	7753	GGAAGCTGATGGCTCCGATGATATCTTTTCAATGTTGAAGAGCGTGTGTGATGAG	7812

Qy	1160	CTCATGATCTCTGGCAAAATTTGGCTGACACCGTTTCGGAGAGGAGTCTTCTCTTCAGGGA	12119
Db	7813	CCTCAGAAAGCAGCGACGCTTGGCTCCGACAGAGTTGGAAAATTCATCTTCCATGCGCTT	78727
Qy	1220	TAGAGAGGATCTTCTGGGCAACAATCTCTCTGGTGT- - -AAACATGAGTTTCTCTC	12757
Db	7873	TGGTCGAGATCTTCAGCTGTGTGTATTCGAGTCTGTGGCAGGACATGGCAACACCC	79323
Qy	1276	ATGATTTGGGACCTATTCAGCTTAATGGAAATNGATCTGGCGGGGTTTGGTCCAGTTT	13353
Db	7933	ATGATTTTCGAGGCTTTCGGGATACGAGAGTTGGGATAGGCGCGAATTTTCGTC- TATTAC	79917
Qy	1336	GAAGCGGGTTTATTTAGATCTCTCCGCTGGTCATCAACGGATATGAGAAATCAGCGG	13959
Db	7992	AACGTGTTGTTTTCACGATCTCTGGACCTGATATCAATGGCTACGAGAGAGACAGCAT	80515
Qy	1396	ATGTGCCCTGAGGAATCTCAGAACTTCACGTCGGATCGCATCGAAGTGGTTAACGCT	14555
Db	8052	CTTTCTATATGGTGGGTTCAACCTTTTGCAAGCTCTGATGGCCATGAAATATTCAGAAA	81111
Qy	1456	GTGTCCTGTGACCCAGCCGATATGCCATTTCAATGTCAGGCGGATTCAGAGAAAACACA	15151
Db	8112	AGCCATGGGAAGACAGCATCTGTTTGTATCCCGCTGTGAATAGCCAGAGGCGGG	81717
Qy	1516	AAATTAAGATTAAGGCTTAAAGAGGGGATATCTGAACCTTAATGATPAAGGTGTGTACA	15757
Db	8172	AGATTGAAGTATGTTGTAACACGGTCTATTCGCTGTTTTCACAGCTCATCATTTGGC	82313
Qy	1576	TCTGACTCGCAAAATATCACTCAGGACTTGGCTGACATCGATACCAATATTTTTCAG	16359
Db	8232	GGCAGTGTGAGGCGCGCTACAGTTGATPAACAGACTGGCGGAGATGAACTTCTTCAGC	82915
Qy	1636	GCACCACTGAACCAAGCGGTGATPAACCAATATGACAGATCGTCAAACTTTCTGTG	16959
Db	8292	TACATATCGAACCGCGCTCGSAAACTCGTCCGCTGTCAATTCAGCACTCTTCATG	83515
Qy	1696	ATGATGACGAAATTTCTGTTAGACCATATCTCCCGCTCTTGTGTCTCATGACGG	17555
Db	8352	GTCAAGAAAGCAAAAGTTTGGTTTAACCTCGGACATCCACAGAGTATGAGACCGATGGG	84111
Qy	1756	ATCCCAAAAGCAGTGTATTTGCTGGACTATGATGCGAGAT	1797
Db	8412	CTTGTCCTGAGCTGTGTTGCATTCAGATCGAATGCCAGCT	8453
RESULT 7			
US-08-459-569-19			
Sequence 19 Application US/08459569			
Patent No. 5543501			
GENERAL INFORMATION:			
APPLICANT: Slighton, Jerry L.			
APPLICANT: Tepfer, David A.			
TITLE OF INVENTION: R1 T-DNA Promoters			
NUMBER OF SEQUENCES: 19			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: GRAY, CARL, AMES & FRYE			
STREET: 401 B Street, Suite 1700			
CITY: San Diego			
STATE: California			
COUNTRY: USA			
ZIP: 92101-4297			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patentin Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08459, 569			
FILING DATE: 02-JUN-1995			
CLASSIFICATION: 530			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US 08/008,216			
FILING DATE: 25-JAN-1993			

APPLICATION NUMBER:	US 06/725,368
FILING DATE:	22-APR-1985
ATTORNEY/AGENT INFORMATION:	
NAME:	Barthorst, Maetrie W.
REGISTRATION NUMBER:	36,740
REFERENCE/DOCKET NUMBER:	P1020US1
TELECOMMUNICATION INFORMATION:	
TELEPHONE:	(619) 699-2700
TELEFAX:	(619) 236-1048
INFORMATION FOR SEQ ID NO:	19:
SEQUENCE CHARACTERISTICS:	
LENGTH:	21126 base pairs
TYPE:	nucleic acid
STRANDEDNESS:	single
TOPOLOGY:	linear
MOLECULE TYPE:	DNA (genomic)
HYPOTHETICAL:	NO
ORGANISM:	Agrobacterium rhizogenes
STRAIN:	STRAIN A4
IMMEDIATE SOURCE:	
LIBRARY:	CONVOLVULUS ARVENSIS PLANT CELLS
CLONE:	CLONE 7
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	complement (937..2262)
OTHER INFORMATION:	//label= ORF1SUBSEQUENCE
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	complement (2649..3458)
OTHER INFORMATION:	//label= ORF2SUBSEQUENCE
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	3726..4799
OTHER INFORMATION:	//label= ORF3SUBSEQUENCE
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	complement (4607..4918)
OTHER INFORMATION:	//label= ORF5SUBSEQUENCE
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	5143..6216
OTHER INFORMATION:	//label= ORF6SUBSEQUENCE
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	complement (5071..5643)
OTHER INFORMATION:	//label= ORF7SUBSEQUENCE
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	6609..8888
OTHER INFORMATION:	//label= ORF8SUBSEQUENCE
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	complement (6576..6830)
OTHER INFORMATION:	//label= ORF9SUBSEQUENCE
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	9748..10044
OTHER INFORMATION:	//label= ORF10SUBSEQUENCE
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	complement (10509..11282)
OTHER INFORMATION:	//label= ORF11SUBSEQUENCE
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	12466..13002
OTHER INFORMATION:	//label= ORF12SUBSEQUENCE
FEATURE:	
NAME/KEY:	misc_feature

LOCATION: 13723..14319
OTHER INFORMATION: /label= ORF13SUBSEQUENC
FEATURE:
NAME/KEY: misc feature
LOCATION: 15659..16210
OTHER INFORMATION: /label= ORF14SUBSEQUENC
FEATURE:
NAME/KEY: misc feature
LOCATION: 16517..17545
OTHER INFORMATION: /label= ORF15SUBSEQUENC
FEATURE:
NAME/KEY: misc feature
LOCATION: 17737..18189
OTHER INFORMATION: /label= ORF16SUBSEQUENC
FEATURE:
NAME/KEY: misc feature
LOCATION: 18177..18743
OTHER INFORMATION: /label= ORF17SUBSEQUENC
FEATURE:
NAME/KEY: misc feature
LOCATION: 19031..19390
OTHER INFORMATION: /label= ORF18SUBSEQUENC
US-08-459-569-19

Query Match 6.8%; Score 121.6; DB 1; Length 21126;
Best Local Similarity 45.7%; Pred. No. 8.4e-32;
Matches 540; Conservative 0; Mismatches 634; Indels 8; Gaps 3;

QY 623 TCGACCTTCTCTACGACTGACAGACCGTTTTCGACCAATGTTCCGATAGTGAACGATCG 662
DB 7273 TTGACCCGCGGTACGACTATCAGGTTTCTCCAGAGTGGAGTCCCATGATGTTGTG 7332
QY 683 GCTTCTTCCGAGAGTGTCTTAAGCGGAAGGGGGTCAATTGGCGCTGGCACTTTCG 742
DB 7333 GGCATTATCCGAGAGGCGCACACTACTAGTGTGTCATGAGGAGGCGGGCTGTCTG 7392
QY 743 GACTCGTGTGGCAAGCACTGCTTCATGCTGGGTAGACATGTTACATATATGAAG 802
DB 7393 GCGTGTGTGTCGACAGAACTACTTGGCGCTGGCGTCAAGAAATCACTTTTCGATA 7452
QY 803 CAATGATCGGT--TGAAGGCAAGCTTTGCTCAATGCTTTTCAGGAGCGTCTTAAG 859
DB 7453 CCGTGTGATGATCCGATGTTTGGGCGATCGCCGATGCCAAAGCGGAGCGTCAACAG 7512
QY 860 TCGTGGCCGAATAGGGGCGATGCGATTTCTCTGTCATCTGCTTTTCTTCC 919
DB 7513 CTTGAGTGTGCTGTGTCATCTTCTCCGCAACCACTTGTCTCTCACTATC 7572
QY 920 TCGAGCGTACGCGCTCTTTCATGAGCGCTTCCCAATCCCGGACAGTGCACACTT 979
DB 7573 TGGATAGTTTAAATTCCTCCAGCTTGTCTTGTGCGGACAGACACACACAG 7632
QY 980 ACTTGTCTACCAAGGCGTCAATATATGAAAGCGGCGAGCTGCCAAGCTGT 1039
DB 7633 CACTATATTTCCGCAAGAAAGCTACGATGACGCGGGGCAAGCTCCCGGGGATAT 7692
QY 1040 TCGATGCGCTTTCACAGGCTTGGGCGCTTCTTGAAGAGCGTTTTCATGAGCGAGATA 1099
DB 7693 TTGAGCGGTGACATGTTGATGAGAACACTACTCTTCAAGAGGTGGAACGGAATGCA 7752
QY 1100 TTGTGTGCTTGCCTGCTGCTATTTACTCAGGCGCTTGAATCAGACACATTAGTGG 1159
DB 7753 GAGACCTGATGCTCCGATGATATCTTTTCATGTTGAAGAGCGTCTGCTGATGAAG 7812
QY 1160 CTGATGCTCTGCAAAATTTGGCTGAACCGTTTCGGAGAGAGTCTTCTTCAAGGA 1219
DB 7813 CCTCAGAGACGCGACCTTGGCTCGAGAGTTCCGAAATTCACCTTTCATGCGCTT 7872
QY 1220 TAAAGAGATCTTTCCTGGGACACATCTCTGTGTGT---AAACATGAGTTTCTC 1275
DB 7873 TGTGTGAGATCTTCACTGTGTGTAATTCAGTCTGTGTGCAAGGACATGCAACACCCC 7932
QY 1276 ATGATTGGAGCTATTCAAGCTAATGGAATAGGATCTGGCGGGGTTTGTTCAGTTT 1335

DB 7933 ATGATTTCAGAGCTTCCGGATCTAGGTTGGGATACGGCCGAGTTTGTCTC-TATTAC 7991
QY 1336 GAAGCGGCTTTATTTAGATCTCCCGCTGGTCTATCAAGGATATGAAGAAATTCAGCG 1395
DB 7992 AACGTGTGTTTTCACAGATCTCGAGCTGATTTATCAATGGCTACAGAGGAGCAGCAT 8051
QY 1396 ATGTGCCCTGAAGGATCTCAGAACTTCCACGCTGGATGCGATCTGAAGTGTTAACGT 1455
DB 8052 CTTTCTATGTGTGGGTTCAACTTTTGCAGGCTCTGATGCGATTAATATTTCCGAAA 8111
QY 1456 GTGTCTGTAGCCAGGACATATGCAATGTTCAAGTCAAGGCGCATTCAGAAAGAAACA 1515
DB 8112 AGCCATGCGAAAGACAGACTCTGTTTGTATCCGCTGCGGAATACCAAGAGGCGGG 8171
QY 1516 AAAATTAAGTAAAGCTTAAAGCGGGAATATCGAACTTATGATTAAGTGTGTGACCA 1575
DB 8172 AGATTGAAGTATGCTTGAACACGCTCATTCGCGTGTTTTGAACGAGTCATCATTTGGC 8231
QY 1576 TCTGACTGCAAAATATCCAACTCAGGCAATTGCTGACATGGGATCAATATTTTCAG 1635
DB 8232 GGCAGTCTGAGGCGGCTCACTGATTAACAGCTGCGGGGATGAGCTTCTTCAGC 8291
QY 1636 GCACCACTGAACCAAGCGTGTGATTAACAGCAATATGACAGATCTCAAAACTTCTC 1695
DB 8292 TACAATATGAACCGCGCTCGGAAACTGTGTCGCGCTGTAATTCAGGACTTTCATG 8351
QY 1696 ATGACTGAACAAATTTCTGTGTTAGACCATATCTCTGCTTGTCTCATGACGG 1755
DB 8352 GTCCACAGCAAAAGTTTGGGTTAATCTCCGATCCAGCACTGATATGAGCCGATGG 8411
QY 1756 ATGCAAAAAGCATGTATTCCTGCTGACTATGATGCGCAGAT 1797
DB 8412 CTGTGCTGAGCTGTGTTGATTCATGCAATGCAATGCCAGCT 8453

RESULT 8
US-08-458-831-19
; Sequence 19, Application US/08458831
; Patent No. 582486
; GENERAL INFORMATION:
; APPLICANT: Slightom, Jerry L.
; APPLICANT: Teflex, David A.
; TITLE OF INVENTION: R1 T-DNA Promoters
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; STREET: 401 B Street, Suite 1700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-4297
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458, 831
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/008, 216
; FILING DATE: 25-JAN-1993
; APPLICATION NUMBER: US 06/725, 368
; FILING DATE: 22-APR-1985
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhorst, Marlene W.
; REGISTRATION NUMBER: 36,740
; REFERENCE/DOCKET NUMBER: P1020US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 699-2700
; TELEFAX: (619) 236-1048

INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 21126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Agrobacterium rhizogenes
STRAIN: STRAIN A4
IMMEDIATE SOURCE:
LIBRARY: CONVOLVULUS ARVENSIS PLANT CELLS
CLONE: CLONE 7
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (937..2262)
OTHER INFORMATION: /label= ORF15SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (2649..3458)
OTHER INFORMATION: /label= ORF2SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 3726..4799
OTHER INFORMATION: /label= ORF3SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (4041..4400)
OTHER INFORMATION: /label= ORF4SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (4607..4918)
OTHER INFORMATION: /label= ORF5SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 5143..6216
OTHER INFORMATION: /label= ORF6SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (5071..5643)
OTHER INFORMATION: /label= ORF7SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 6609..8888
OTHER INFORMATION: /label= ORF8SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (6576..6830)
OTHER INFORMATION: /label= ORF9SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 9748..10044
OTHER INFORMATION: /label= ORF10SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (10509..11282)
OTHER INFORMATION: /label= ORF11SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 12465..13002
OTHER INFORMATION: /label= ORF12SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 13723..14319
OTHER INFORMATION: /label= ORF13SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 15655..16210
OTHER INFORMATION: /label= ORF14SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (16517..17545)

OTHER INFORMATION: /label= ORF15SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (17737..18189)
OTHER INFORMATION: /label= ORF16SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (18177..18743)
OTHER INFORMATION: /label= ORF17SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (19031..19390)
OTHER INFORMATION: /label= ORF18SUBSEQUENCE
US-08-458-831-19
Query Match 6.8%; Score 121.6; DB 1; Length 21126;
Best Local Similarity 45.7%; Pred. No. 8.4e-32;
Matches 540; Conservative 0; Mismatches 634; Indels 8; Gaps 3;
Qy 623 TCGACTTGGCTCTACGACTGCGAGACCGTTTGTGACCAATGTTCCGATAGTGAAGATCG 682
Db 7273 TTACCTGGGGTATCGACTATACAGTTTCTCTCCAGAGTGCATGCCATGATGATGTGG 7332
Qy 683 GCTTCTTTCCGAGAGTGTCTTAAGCCGAAGTGGCGGTCATTGGCCCTGSCATTTCG 742
Db 7333 GGCATTATCCGGAAGCGCACCACTACTAGTGTCCATCGTAGGAGGCGGCTGTCTG 7392
Qy 743 GACTCGTGGGCGAAGCAAGATGCTTCATGCTGGGGTAGACGATGTACAAATATATGAAG 802
Db 7393 GCCTTGTTGTGCAAGAACTACTAGTGGCGTGGCGTCAAGAAATACCTCTTTTGATA 7452
Qy 803 CAAGTATCGTGT---TGAGGCAAGCTTGTGTCATGCTTTTCAGGAGCGTCTAGTG 859
Db 7453 CGCTTATAGATCGGTAGTTTGGGGCATGCCATGCCAAGCGGAGCGTCCACAGG 7512
Qy 860 TCGTGGCCGAATGGGGCGATCCGATTTCTCTGCTGATCTGTTTCTTTTCTTC 919
Db 7513 CTTTGACGTGCTGCGGTGTCATGCTTTCGCCCAACCACTTTCGTGTCTATCTATC 7572
Qy 920 TCGAGGTTACGGCGTGTCTTCGATGAGGCGTTCCTCCAAATCCCGACAGTGCACCTT 979
Db 7573 TGGATTAATTTAGAAATTCCTCCAGCTTCCTTCTGTGCTGGCAAGACACACACAG 7632
Qy 980 ACTTGGTCTACCAAGGCGTCCAAATATGATGGAAGCCGGGAGCTGCCACGAAGCTGT 1039
Db 7633 CACTATATTTCCGCCAAGAAACGCTACGCAATGCAACGGGGGCAAGCTCCGGGGATAT 7692
Qy 1040 TCCATGCGCTTACAAAGCTTGGCGTTCCTTGAAGACGCTTTTCATGACGAGATA 1099
Db 7693 TTACGGGGGTACATGTGCGATGGAAGCACTACTTACCAAGGCTGTGAACGGAATGGCA 7752
Qy 1100 TTGTGTGGCTTGGCGGTGCTATTAATCAAGGCTTTGAAATACAGACACATTAGTGGG 1159
Db 7753 GGAAGCTGATGCTCGATGATGATATCTTTTCAATGTAAGAGCGCTGCTGATGATAG 7812
Qy 1160 CTCATGATCTCTGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTTCACAGGA 1219
Db 7813 CTCACAGAGACGGACGCTTTGGCTCCAGAGTTCGGAATAATCACTTTCATGCGGTTT 7872
Qy 1220 TAAGAGAGATCTTTCGGGCAACATCTCTCTGTGT---AAACATGAGATTTCTTC 1275
Db 7873 TGTGCGAGATCTTCAGCTGTGTAAATTCGAGTCTGTGCAAGGCAAGCAACACCCC 7932
Qy 1276 ATGATTGGACCTATTCACACTAATGGAATAGAGTCTGGGGGTTGTGTCAGTCTT 1335
Db 7933 ATGATTTCGAGGCTTTCGGATCTAGAGTTGGAATAGGCGCAAGTTTCGTC-TATTAC 7991
Qy 1336 GAAAGCGGTTTATAGATCTCTCGCTTGTGATCAACGATATGAAGAAATCAAGCGG 1395
Db 7992 AACGTGTGTTTCAACGATCTGAGCTGATATATCAATGGCTACAGAGGACACACAT 8051
Qy 1396 ATGTGCTGAGGAATCTCAGAACTTCAACGTCGATCGCATCTGAAGTGTGTTACGCT 1455

Db 8052 CTTTCTATGTTGGGGTTCAACTTTTGCAGGCTCTGATGCGCATTTGAATAATTTCCAGAAA 8111
Qy 1456 GTGCTGTGAGCCAGCCATATGCGCATTTCAAGTTCAGGGGATTGAGAAAGAAAACA 1515
Db 8112 AGCCATCGAAGAGACCACTCTGTTTGTATCCGCTGGTGAAGCCAGAGAGGGCGG 8171
Qy 1516 AAAATAAAGTAAGCTTAAAGAGCGGATATCTGAACTTTATGATTAAGTGTGTGACA 1575
Db 8172 AGATTGAAGGATCTTGAAACAGCGTCATTCGGGTGTTTTTTCAGAGGTCATCATTTGGC 8231
Qy 1576 TCTGAGATCCGAATATTCACACTAGGCATTTGCTGACATGCGATACCAATATTTTCAG 1635
Db 8232 GCGAGTCTAGGCGCGCTAGCATGTATGACAGCTGCGGGGATGAGACTTCTTCAGC 8291
Qy 1636 GCACAGTAGAACCAGGCGTTGATATACAGCATATGACAGATCGTAAACTTCTCTG 1695
Db 8292 TACATATCGAACCCCGCTGGAACTCTGCTGCGCTGCAATTAGCATCTTTCATG 8351
Qy 1696 ATGACTGAGCAAAATTTCTGTTAGACCATATCTCCGCTTGTGCTCATGAGACGGG 1755
Db 8352 GTACGAGAGAAAAGTTTGGGTAACTCGGGCATCCAGAGATGATAGACGAGTGG 8411
Qy 1756 ATGCAAAAGACAGTGTATGCTGCTGACTATGATGATGCGAGAT 1797
Db 8412 CTGTCCGTAGAGTGTGTCATTCATGACATCGAATCGCACACT 8453

RESULT 9

US-09-230-388-2
; Sequence 2, Application US/09230388
; Patent No. 6291644
; GENERAL INFORMATION:
; APPLICANT: Iwamoto, Mitsunori; Jung, Sang-Kee
; TITLE OF INVENTION: NOVEL PROTEIN, ITS GENE, REAGENTS FOR INDUCING
; FILE REFERENCE: 4703/0214
; CURRENT APPLICATION NUMBER: US/09/230,388
; CURRENT FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: PCT/J998/02261
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1575
; TYPE: DNA
; ORGANISM: Scomber japonicus
US-09-230-388-2

Query Match 2.1%; Score 38.6; DB 3; Length 1575;
Best Local Similarity 66.1%; Pred. No. 0.0092;
Matches 72; Conservative 0; Mismatches 34; Indels 3; Gaps 1;

Qy 713 AAGTGGCGTCAATGCGCTGCGCATTTCCGACTCGTGTGCGAAAGCAAGTCTTCATG 772
Db 182 ATGTGTTATAGTCGAGCTGCGATGCGGACGAGCGGCGCAAGTTACTGCAAGAG 241
Qy 773 CTGGCGTAGCGATGTTACATATATATGAAGCAAGTATGCTGTTGAGG 821
Db 242 CAGG---ACACACGGTAACCATATGTGAGGCTATGATGCTGTTGAGG 287

RESULT 10

US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00

; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 2.1%; Score 37; DB 3; Length 4403765;
Best Local Similarity 57.3%; Pred. No. 16;
Matches 67; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 709 CCGAAGTGGCGGTCAATGCGCTGCGCATTTCCGAGCTGCTGTGCGCAAGCAAGTCTT 768
Db 993867 CCGACGTTGCCGTGCTGCGAGCGGTAATGTCGGAATGTGCTCGCAATTACGTTGCTG 993926
Qy 769 CATGCTGGGTAGACAGATGTTACATATATATGAAGCAAGTATGCTGTGAGAGGCAAG 825
Db 993927 AGCGCAGGATTAATGATGCTGTCATCTATGAAAGAGCCGACGATGTTGGCGGAAG 993983

RESULT 11

US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 2.1%; Score 37; DB 3; Length 4411529;
Best Local Similarity 57.3%; Pred. No. 16;
Matches 67; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 709 CCGAAGTGGCGGTCAATGCGCTGCGCATTTCCGACTGCTGTGCGCAAGCAAGTCTT 768
Db 993866 CCGACGTTGCCGTGCTGCGAGCGGTAATGTCGGAATGTGCTCGCAATTACGTTGCTG 993925
Qy 769 CATGCTGGGTAGACAGATGTTACATATATATGAAGCAAGTATGCTGTGAGAGGCAAG 825
Db 993926 AGCGCAGGATTAATGATGCTGTCATCTATGAAAGAGCCGACGATGTTGGCGGAAG 993982

RESULT 12

US-09-415-277C-4/C
; Sequence 4, Application US/09415277C
; Patent No. 6531308
; GENERAL INFORMATION:
; APPLICANT: Hershenberger, Charles
; APPLICANT: Payson, Robert
; TITLE OF INVENTION: Ketoreductase Gene and Protein from Yeast
; FILE REFERENCE: X-11325A
; CURRENT APPLICATION NUMBER: US/09/415,277C

/ CURRENT FILING DATE: 1999-10-08
/ PRIOR APPLICATION NUMBER: US 09/182,985
/ PRIOR FILING DATE: 1998-10-30
/ NUMBER OF SEQ ID NOS: 17
/ SOFTWARE: PatentIn version 3.1.
/ SEQ ID NO 4
/ LENGTH: 1032
/ TYPE: DNA
/ ORGANISM: s. cerevisiae
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1032)
/ OTHER INFORMATION:
US-09-415-277C-4

Query Match 2.0%; Score 35.2; DB 4; Length 1032;

Best Local Similarity 53.7%; Pred. No. 0.11;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 1498 ATTCAGAAAGGAAAGACAAATAAAGATTAAGGCTTAAGCGGATATCTGAATTTAT 1557
Db 624 ATCGAAAGTTGAGACCAAAAGAGAGAGAGGTTGACTGTTGTTGATTTGAT 565
Qy 1558 GATTAAGTGTGTGTCACATCTGCACTGCAAAATATCCAACTCAGGCAATTCCTGACATGC 1617
Db 564 GTGATCTTCATCTCTTTTGTGAACCTCCAGGAGCCTTTTCAGCAAACTTCTTGATGC 505
Qy 1618 GATACCAATATTTTTC 1633
Db 504 AAAGTAAGCATTTATC 489

RESULT 13

US-09-415-277C-6/c
/ Sequence 6, Application US/09415277C
/ Patent No. 6531308
/ GENERAL INFORMATION:
/ APPLICANT: Hershenberger, Charles
/ APPLICANT: Payson, Robert
/ TITLE OF INVENTION: Ketoreductase Gene and Protein from Yeast
/ FILE REFERENCE: X-11325A
/ CURRENT APPLICATION NUMBER: US/09/415, 277C
/ CURRENT FILING DATE: 1999-10-08
/ PRIOR APPLICATION NUMBER: US 09/182,985
/ PRIOR FILING DATE: 1998-10-30
/ NUMBER OF SEQ ID NOS: 17
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 6
/ LENGTH: 1032
/ TYPE: RNA
/ ORGANISM: s. cerevisiae
US-09-415-277C-6

Query Match 2.0%; Score 35.2; DB 4; Length 1032;

Best Local Similarity 53.7%; Pred. No. 0.11;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 1498 ATTCAGAAAGGAAAGACAAATAAAGATTAAGGCTTAAGCGGATATCTGAATTTAT 1557
Db 624 ATCGAAAGTTGAGACCAAAAGAGAGAGAGGTTGACTGTTGTTGATTTGAT 565
Qy 1558 GATTAAGTGTGTGTCACATCTGCACTGCAAAATATCCAACTCAGGCAATTCCTGACATGC 1617
Db 564 GTGATCTTCATCTCTTTTGTGAACCTCCAGGAGCCTTTTCAGCAAACTTCTTGATGC 505
Qy 1618 GATACCAATATTTTTC 1633
Db 504 AAAGTAAGCATTTATC 489

/ Patent No. 6387658
/ GENERAL INFORMATION:
/ APPLICANT: Luo, Ying
/ APPLICANT: Huang, Betty
/ TITLE OF INVENTION: No. 6387658el PCNA-Associated Cell Cycle Proteins, Compositions
/ FILE REFERENCE: A68291/DJB/RMS/DAV
/ CURRENT APPLICATION NUMBER: US/09/404,627
/ CURRENT FILING DATE: 1999-09-23
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 3
/ LENGTH: 3593
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-404-627-3

Query Match 2.0%; Score 35.2; DB 4; Length 3593;

Best Local Similarity 57.1%; Pred. No. 0.29;
Matches 64; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 1503 GAAGGAAAGACAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAATTTATGATA 1562
Db 2397 GAGAAATAAATTAAATACAGATTAACATTCATTTACAGATGTCAAGATCAGAA 2456
Qy 1563 GGTGTGTGTGTCACATCTGCACTGCAAAATATCCAACTCAGGCAATTCCTGACATGC 1614
Db 2457 GATAGTATGTCTATCTGTCTTACAAATACACAACTGAGATTTCCCGACA 2508

RESULT 15

US-09-404-627-1
/ Sequence 1, Application US/09404627
/ Patent No. 6387658
/ GENERAL INFORMATION:
/ APPLICANT: Luo, Ying
/ APPLICANT: Huang, Betty
/ TITLE OF INVENTION: No. 6387658el PCNA-Associated Cell Cycle Proteins, Compositions
/ FILE REFERENCE: A68291/DJB/RMS/DAV
/ CURRENT APPLICATION NUMBER: US/09/404,627
/ CURRENT FILING DATE: 1999-09-23
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 4205
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-404-627-1

Query Match 2.0%; Score 35.2; DB 4; Length 4205;

Best Local Similarity 57.1%; Pred. No. 0.33;
Matches 64; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 1503 GAAGGAAAGACAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAATTTATGATA 1562
Db 3009 GAGAAATAAATTAAATACAGATTAACATTCATTTTACAGATGTCAAGATCAGAA 3068
Qy 1563 GGTGTGTGTGTCACATCTGCACTGCAAAATATCCAACTCAGGCAATTCCTGACATGC 1614
Db 3069 GATAGTATGTCTATCTGTCTTACAAATACACAACTGAGATTTCCCGACA 3120

Search completed: November 23, 2003, 21:27:11
Job time: 119.342 secs

RESULT 14
US-09-404-627-3
/ Sequence 3, Application US/09404627

181 GCTTGAAGAGGCTGCGCATGTCGCTCCCGAGATCTAGCTGGGAAAGTAGCA 240
181 GCTTGAAGAGGCTGCGCATGTCGCTCCCGAGATCTAGCTGGGAAAGTAGCA 240
241 GTTCTCTCCGCTTATCTATATTTGGCAAGAAATTCGGGGCGGATCTTGAATGAAA 300
241 GTTCTCTCCGCTTATCTATATTTGGCAAGAAATTCGGGGCGGATCTTGAATGAAA 300
301 CTTTGGGGGGGGGCAAGAGTAGTGTCTGTTGGCCATTCGACTTGGACACATTTGGATG 360
301 CTTTGGGGGGGGGCAAGAGTAGTGTCTGTTGGCCATTCGACTTGGACACATTTGGATG 360
361 GATTTCCTCCAGACACAATAATCCAAAGCCCTGTTTGGCTGAGCGGTAAAGATGCA 420
361 GATTTCCTCCAGACACAATAATCCAAAGCCCTGTTTGGCTGAGCGGTAAAGATGCA 420
421 CCGATGATCTTATGTCATTTGTCGTCATTTCAATCTTAAGACTGCGCGCTTTCGACC 480
421 CCGATGATCTTATGTCATTTGTCGTCATTTCAATCTTAAGACTGCGCGCTTTCGACC 480
481 CTGCGCATGCGCGCTTACGAGAAATGAGACAGATGAAGCCGTTACCGGGTTTACCATACC 540
481 CTGCGCATGCGCGCTTACGAGAAATGAGACAGATGAAGCCGTTACCGGGTTTACCATACC 540
541 CTTGAAGAGGCGCGTCCGCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 600
541 CTTGAAGAGGCGCGTCCGCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 600
601 TCGGCAAGGTTCTTTCCCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 660
601 TCGGCAAGGTTCTTTCCCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 660
661 TGTTCGATAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
661 TGTTCGATAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
721 GTGATGCGCGCTGATTTCCGAGCTGATGATGATGATGATGATGATGATGATGATGATG 780
721 GTGATGCGCGCTGATTTCCGAGCTGATGATGATGATGATGATGATGATGATGATGATG 780
781 GAGCATTTTACATATATGAAAGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
781 GAGCATTTTACATATATGAAAGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
841 TTTAAGGAGCGCTCTAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 900
841 TTTAAGGAGCGCTCTAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 900
901 TTTGCTGTTGTTTCTTCTGAGCGCTTACGAGCGCTTACGAGCGCTTACGAGCGCTT 960
901 TTTGCTGTTGTTTCTTCTGAGCGCTTACGAGCGCTTACGAGCGCTTACGAGCGCTT 960
961 CCGGCGCAGTGCACCTTACTTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1020
961 CCGGCGCAGTGCACCTTACTTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1020
1021 CAGTGCACCGAAGCGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
1021 CAGTGCACCGAAGCGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
1081 GGTTCATGAGCAGATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
1081 GGTTCATGAGCAGATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
1141 TCAAGACATTAAGTGGGCTCATGATCTCGGCAATTTGGCTGGAACCGTTTGGGAGG 1200
1141 TCAAGACATTAAGTGGGCTCATGATCTCGGCAATTTGGCTGGAACCGTTTGGGAGG 1200
1201 GAGTCCCTTCTTCAAGGATGAGAGATCTTTTGGGCAACATCTCTCGGAGGTTAAA 1260
1201 GAGTCCCTTCTTCAAGGATGAGAGATCTTTTGGGCAACATCTCTCGGAGGTTAAA 1260
1261 CATGAGTTTCTCTCATGATTTGGGACTATTTCAAGCTAATGGGAATAGATCTTGGCGGG 1320

1261 CATGAGTTTCTCTCATGATTTGGGACTATTTCAAGCTAATGGGAATAGATCTTGGCGGG 1320
1321 TTTGTCACGTTTGTGAAAGCGGTTTATGAAATCTCCGCTTGGTATCAACGATAT 1380
1321 TTTGTCACGTTTGTGAAAGCGGTTTATGAAATCTCCGCTTGGTATCAACGATAT 1380
1381 GAAGAAATGAGGAGATGTCCTGTAAGGAATCTCAAGAACTTCACGCGGATGCACT 1440
1381 GAAGAAATGAGGAGATGTCCTGTAAGGAATCTCAAGAACTTCACGCGGATGCACT 1440
1441 GAAGGTTTAAAGGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1500
1441 GAAGGTTTAAAGGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1500
1501 CAGAGGAAAGCAAAATTAAGATTAAGGCTTAAAGCGGATATCTGAATCTTATGAT 1560
1501 CAGAGGAAAGCAAAATTAAGATTAAGGCTTAAAGCGGATATCTGAATCTTATGAT 1560
1561 AAGTGGTGTCAATCTGAGCTGCGCAATATCACTCAGGATTTGCTGACATGCGAT 1620
1561 AAGTGGTGTCAATCTGAGCTGCGCAATATCACTCAGGATTTGCTGACATGCGAT 1620
1621 ACCAATATTTTTCAGGACACAGTGAACCAAGCGTTGATTAACAGCATATGACAGATCG 1680
1621 ACCAATATTTTTCAGGACACAGTGAACCAAGCGTTGATTAACAGCATATGACAGATCG 1680
1681 TCAAAATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
1681 TCAAAATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
1741 GTCTCATGAGACGGGATCGCAAAAGCAGTGTATTGCTGAGATGATGATGATGATGAT 1800
1741 GTCTCATGAGACGGGATCGCAAAAGCAGTGTATTGCTGAGATGATGATGATGATGATG 1800
1801 A 1801
1801 A 1801
RESULT 2
US-10-187-339-7
; Sequence 7, Application US/10187339
; Publication No. US20030084478A1
; GENERAL INFORMATION:
; APPLICANT: Ream, Malt et al.,
; TITLE OF INVENTION: Plants Having Enhanced Gall Resistance and
; FILE REFERENCE: Methods and Compositions for Producing the Same
; CURRENT APPLICATION NUMBER: US/10/187,339
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US/09/434,837
; PRIOR FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 60/107,185
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO: 7
; LENGTH: 1807
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-187-339-7
Query Match 99.3%; Score 1788; DB 14; Length 1807;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
12 ACCTTCCTTGAATTAACAGTGCATATCTCCCAACCAAAATGATGATGATGATGATGATGAT 71
18 ACCTTCCTTGAATTAACAGTGCATATCTCCCAACCAAAATGATGATGATGATGATGATGAT 77
72 CGATTAAGGCGGATGAATTTGGACCGCAGGCTTTCGATGCTTCTTAAGAGAGAGCTTC 131

Db 78 CGATTAAGCGGATGAATTGACCGCAGGTTTCCGATGCTTCTTAGAAGAGAGCTTC 137
Qy 132 TAGGGAGAGAGGATTAATCAAGCTCCACGAGTGAAGCGGTGGTGTAGCTTGGCAAAG 191
Db 138 TAGGGAGAGAGGATTAATCAAGCTCCACGAGTGAAGCGGTGGTGTAGCTTGGCAAAG 197
Qy 192 GCTGGCCGATGCTGCTTCCCGAGATCTCAGCTGTGGAAGGTAGCAGTCTCTCCG 251
Db 198 GCTGGCCGATGCTGCTTCCCGAGATCTCAGCTGTGGAAGGTAGCAGTCTCTCCG 257
Qy 252 TTATATTAATTTGGCAAGAAATTCGGGGCGGATCTTGAATCGAACTTGGGGCG 311
Db 258 TTATATTAATTTGGCAAGAAATTCGGGGCGGATCTTGAATCGAACTTGGGGCG 317
Qy 312 GGCAACGATGAGTGTGCTGCTGCGATCGACTTGGACACCAATTTGCAATTTCTCCGA 371
Db 318 GGCAACGATGAGTGTGCTGCTGCGATCGACTTGGACACCAATTTGCAATTTCTCCGA 377
Qy 372 AGCAACAATAATCCAGCCCTGTTTTGCTGAGCGGTAAAGATGTGCAACCGATGATCT 431
Db 378 AGCAACAATAATCCAGCCCTGTTTTGCTGAGCGGTAAAGATGTGCAACCGATGATCT 437
Qy 432 TAGTCAATTTGTGGCCATTTCATCTCTAAGCTGCCGCTTTGAACTCCCAATGCC 491
Db 438 TAGTCAATTTGTGGCCATTTCATCTCTAAGCTGCCGCTTTGAACTCCCAATGCC 497
Qy 492 GCTGTACGAATAGCGACGATGAAGTGCCTTACCGGGTTTACCAATACCTTGAAGGGC 551
Db 498 GCTGTACGAATAGCGACGATGAAGTGCCTTACCGGGTTTACCAATACCTTGAAGGGC 557
Qy 552 CGTGCCATTGACATGCTTGTAGCTTGTGCAAACTGATGCTGAAAGGTTGCGCAGGTT 611
Db 558 CGTGCCATTGACATGCTTGTAGCTTGTGCAAACTGATGCTGAAAGGTTGCGCAGGTT 617
Qy 612 CTTTCCAAATCGACTGCTCTAAGACTGCAAGCCGTTTTTGAACCAATGTTCCGATAG 671
Db 618 CTTTCCAAATCGACTGCTCTAAGACTGCAAGCCGTTTTTGAACCAATGTTCCGATAG 677
Qy 672 TGAACGATGCGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCGTCAATGGCCG 731
Db 678 TGAACGATGCGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCGTCAATGGCCG 737
Qy 732 TGCAATTTCCGAGCTCGTGGTGGCAACGAATGCTTCAATGCTGGGTTAGACATGTTAC 791
Db 738 TGCAATTTCCGAGCTCGTGGTGGCAACGAATGCTTCAATGCTGGGTTAGACATGTTAC 797
Qy 792 AATATATGAAGCAAGTATGCTGTGGAGGCAAGCTTGTGCAATGCTTTCAGGAGCGC 851
Db 798 AATATATGAAGCAAGTATGCTGTGGAGGCAAGCTTGTGCAATGCTTTCAGGAGCGC 857
Qy 852 TCTTAGTGTGCTGGCCGAATGGGGGCGATGCGATTTCTCTGCTGCAATTTGCTTGT 911
Db 858 TCTTAGTGTGCTGGCCGAATGGGGGCGATGCGATTTCTCTGCTGCAATTTGCTTGT 917
Qy 912 TTTCTTCTCGAGGTTAGGCGCTGCTTGCATGAGGCGCGTCCCAATCCCGGACAGT 971
Db 918 TTTCTTCTCGAGGTTAGGCGCTGCTTGCATGAGGCGCGTCCCAATCCCGGACAGT 977
Qy 972 CGACACTTACTTGTGTACCAAGCGCTCAATACATGTGAAAAGCCGGGACAGTGCACC 1031
Db 978 CGACACTTACTTGTGTGTACCAAGCGCTCAATACATGTGAAAAGCCGGGACAGTGCACC 1037
Qy 1032 GAAGCTGTTCATCGCTTTTACAACGCTTGGCGTGTCTTTGAAGAGCGTTTTCATGA 1091
Db 1038 GAAGCTGTTCATCGCTTTTACAACGCTTGGCGTGTCTTTGAAGAGCGTTTTCATGA 1097
Qy 1092 GCGAGATATTTGTGGCTTCCGCTGCGTATTTACTCAGGCTTGAATCAGAGACAT 1151
Db 1098 GCGAGATATTTGTGGCTTCCGCTGCGTATTTACTCAGGCTTGAATCAGAGACAT 1157
Qy 1152 TAGGTGGGCTCATGACTCTGTGGCAATTTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1211
Db 1158 TAGGTGGGCTCATGACTCTGTGGCAATTTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1217

Qy 1212 TTCAGGATATAGAGGATCTTTCTGGGACACATCTCCGNGGTGTAACATGAGATTT 1271
Db 1218 TTCAGGATATAGAGGATCTTTCTGGGACACATCTCCGNGGTGTAACATGAGATTT 1277
Qy 1272 CCTCATGATTTGGGACCTTATCAAGCTAATGGGAATGAGATCTGGCGGGTTTGGTCAAT 1331
Db 1278 CCTCATGATTTGGGACCTTATCAAGCTAATGGGAATGAGATCTGGCGGGTTTGGTCAAT 1337
Qy 1332 TTTTGAAGCGGGTTTATGAGATCTCTCCGTTGGTTCATCAACGGAATAGAAATCA 1391
Db 1338 TTTTGAAGCGGGTTTATGAGATCTCTCCGTTGGTTCATCAACGGAATAGAAATCA 1397
Qy 1392 GCGAGTGCCTCCGAAGGATCTCAGAACTTCAGCTCGGATGCGATCTGAATGGTTAA 1451
Db 1398 GCGAGTGCCTCCGAAGGATCTCAGAACTTCAGCTCGGATGCGATCTGAATGGTTAA 1457
Qy 1452 CGGTGTCTGTGAGCGCAGCGCATATGCGATGTTCAAGTCAGGGCGATTCAAGAGAAA 1511
Db 1458 CGGTGTCTGTGAGCGCAGCGCATATGCGATGTTCAAGTCAGGGCGATTCAAGAGAAA 1517
Qy 1512 GACAAAATTAAGATTAAGCTTAAAGCGGATATTTGAACCTTTATGATTAAGTGGT 1571
Db 1518 GACAAAATTAAGATTAAGCTTAAAGCGGATATTTGAACCTTTATGATTAAGTGGT 1577
Qy 1572 CACATCTGACTCGCAATATCAACTCAAGGATTTGCCGACATGCGATACCAATATTTT 1631
Db 1578 CACATCTGACTCGCAATATCAACTCAAGGATTTGCCGACATGCGATACCAATATTTT 1637
Qy 1632 TCAGGACCAAGTAACCAAGCGGTGATTAACAGCCATATGACAGATTCGTCAAACTCT 1691
Db 1638 TCAGGACCAAGTAACCAAGCGGTGATTAACAGCCATATGACAGATTCGTCAAACTCT 1697
Qy 1692 CCGATGACTGAACGAAATTTCTGTTAACAATATCTCCGTTGTGTCTCATGGA 1751
Db 1698 CCGATGACTGAACGAAATTTCTGTTAACAATATCTCCGTTGTGTCTCATGGA 1757
Qy 1752 CGGATCGCAAAAAGCATATTTCCCTGGAATATGATGCGAGATCC 1799
Db 1758 CGGATCGCAAAAAGCATATTTCCCTGGAATATGATGCGAGATCC 1805

RESULT 3

US-10-187-339-9
; Sequence 9, Application US/10187339
; Publication No. US20030084478A1
; GENERAL INFORMATION:
; APPLICANT: Ream, Malt et al.,
; TITLE OF INVENTION: Plants Having Enhanced Gall Resistance and
; FILE REFERENCE: Methods and Compositions for Producing the Same
; CURRENT APPLICATION NUMBER: US/10/187,339
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US/09/434,837
; PRIOR FILING DATE: 1999-11-04
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2544
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
; US-10-187-339-9

Query Match 99.3%; Score 1788; DB 14; Length 2544;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12 ACCTCTCTTGAATTAACGATGCGATCATCTCCCAACCAAAATGTTGATCTGACAAATGAT 71
Db 318 ACCTCTCTTGAATTAACGATGCGATCATCTCCCAACCAAAATGTTGATCTGACAAATGAT 377

QY 72 CGATAAGGCGGATGATGATGAGACCGCAGGGTTTCGATGCTCTTCTTAGAAGAGAGCTTC 131
DB 378 CGATTAAGCGGATGATGATGAGACCGCAGGGTTTCGATGCTCTTCTTAGAAGAGAGCTTC 437
QY 132 TAGGGGAGGAGGATTAATCTCAAGCTCCACCGAGTGCAGCGCTGGGTTAGCTTGCAGAAAG 191
DB 438 TAGGGGAGGAGGATTAATCTCAAGCTCCACCGAGTGCAGCGCTGGGTTAGCTTGCAGAAAG 497
QY 192 GCTGGCCGATGCTGCTTCCCGAGATCTGAGCTGGTGAAGATGAGTACGATCTCTCCG 251
DB 498 GCTGGCCGATGCTGCTTCCCGAGATCTGAGCTGGTGAAGATGAGTACGATCTCTCCG 557
QY 252 TTATATCTATATTGGCAAGAAATTTCTGGGGCGGATCTTGAAATCGAAACCTTGGGGCG 311
DB 558 TTATATCTATATTGGCAAGAAATTTCTGGGGCGGATCTTGAAATCGAAACCTTGGGGCG 617
QY 312 GGGACAGTGAAGTGTCTGCTTGCATCGACTTGGACACCAATTTGACATGATTTCTCGA 371
DB 618 GGGACAGTGAAGTGTCTGCTTGCATCGACTTGGACACCAATTTGACATGATTTCTCGA 677
QY 372 AGGACAACTAATCCAGCCCTGTTTTGCTGAGGGGTTAAAGATGTCACCGATGATCT 431
DB 678 AGGACAACTAATCCAGCCCTGTTTTGCTGAGGGGTTAAAGATGTCACCGATGATCT 737
QY 432 TAGCATTTTGTGGCCATTTCAATCTCTAAGACTGCGGCTTTCGAAACCTTGCATGCGC 491
DB 738 TAGCATTTTGTGGCCATTTCAATCTCTAAGACTGCGGCTTTCGAAACCTTGCATGCGC 797
QY 492 GCTGTACGAGATGAGCAGATGAAGATCGTTACCGGGTTTACATTAACCTTGAAGGGGC 551
DB 798 GCTGTACGAGATGAGCAGATGAAGATCGTTACCGGGTTTACATTAACCTTGAAGGGGC 857
QY 552 CGTGCCATTTGACATGCTTACCTTAATGCTGAAGATGCTGAAGGTTTCGAGAGTTTC 611
DB 858 CGTGCCATTTGACATGCTTACCTTAATGCTGAAGATGCTGAAGGTTTCGAGAGTTTC 917
QY 612 CTTCCTCAACATGCACTTGTCTCTACGACTGAGACCGTTTTTTTACCAATGTTCCGATAG 671
DB 918 CTTCCTCAACATGCACTTGTCTCTACGACTGAGACCGTTTTTTTACCAATGTTCCGATAG 977
QY 672 TGAACGATCGGCTCTTCTTCGAGAGATGTTCTTAACCGGAAATGCGGCTATTTGCGC 731
DB 978 TGAACGATCGGCTCTTCTTCGAGAGATGTTCTTAACCGGAAATGCGGCTATTTGCGC 1037
QY 732 TGGGATTTCCGGGACTCGTGGTGGCAACGAACTGCTTCACTGCTGGGTGAACGATGTTAC 791
DB 1038 TGGGATTTCCGGGACTCGTGGTGGCAACGAACTGCTTCACTGCTGGGTGAACGATGTTAC 1097
QY 792 AATATATGAAGCAAGTATCGTGTGGAGGCAAGCTTGTGCATGCTTTCAGAGAGCGC 851
DB 1098 AATATATGAAGCAAGTATCGTGTGGAGGCAAGCTTGTGCATGCTTTCAGAGAGCGC 1157
QY 852 TCCTAGTGTGCTGCGCGGAAATGGGGGCGATGCGATTTCTCTGCTGCTGCTTCTGCT 911
DB 1158 TCCTAGTGTGCTGCGCGGAAATGGGGGCGATGCGATTTCTCTGCTGCTGCTTCTGCT 1217
QY 912 TTTTCTTCTGAGAGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
DB 1218 TTTTCTTCTGAGAGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1277
QY 972 CGACACTTACTGTGCTTACCAAGGCTGCAATATACATGATGAAAGCGGGGAGCTGCCACC 1031
DB 1278 CGACACTTACTGTGCTTACCAAGGCTGCAATATACATGATGAAAGCGGGGAGCTGCCACC 1337
QY 1032 GAAGCTTTCATCGCGCTTACCAAGGCTGCGCTTCTTGAAGACGCGTTTTCATGA 1091
DB 1338 GAAGCTTTCATCGCGCTTACCAAGGCTGCGCTTCTTGAAGACGCGTTTTCATGA 1397
QY 1092 GCGAGATATTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1151
DB 1398 GCGAGATATTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1457
QY 1152 TAGGTGGGCTCATGATCTGCTGCGCAAAATTTGGCTGAACGTTTGGAGGAGTCTTCTC 1211

DB 1458 TAGGTGGGCTCATGATCTGCTGCGCAAAATTTGGCTGAACCGTTTGGAGGAGTCTTCTC 1517
QY 1212 TTCAAGGATGAGAGATCTTCTTGGGACACATCTCTCTGCTGCTGCTGCTGCTGCT 1271
DB 1518 TTCAAGGATGAGAGATCTTCTTGGGACACATCTCTCTGCTGCTGCTGCTGCTGCT 1577
QY 1272 CCTCATGATTGGGACCTAATTCAGCTAATGGGAATGAGATCTGGCGGGTTTGGTCCAGT 1331
DB 1578 CCTCATGATTGGGACCTAATTCAGCTAATGGGAATGAGATCTGGCGGGTTTGGTCCAGT 1637
QY 1332 TTTTGAAGCGGGTTTATTGAGATCTCTGCTTGTCTATCAACGATATGAAGAAATCA 1391
DB 1638 TTTTGAAGCGGGTTTATTGAGATCTCTGCTTGTCTATCAACGATATGAAGAAATCA 1697
QY 1392 GCGGATGCTGCTTGAAGGAATCTCAGACTTCAAGCTGCTGCTGCTGCTGCTGCTGCT 1451
DB 1698 GCGGATGCTGCTTGAAGGAATCTCAGACTTCAAGCTGCTGCTGCTGCTGCTGCTGCT 1757
QY 1452 CGGTGTGCTGTGAGCGAGCGCATATGCCATGTTCAAGTCAGGGCGATTGAGAGGAAA 1511
DB 1758 CGGTGTGCTGTGAGCGAGCGCATATGCCATGTTCAAGTCAGGGCGATTGAGAGGAAA 1817
QY 1512 GACAAAATTAAGATTAAGGCTTAAGACCGGATATCTGAACCTTATGATTAAGGTGCTGT 1571
DB 1818 GACAAAATTAAGATTAAGGCTTAAGACCGGATATCTGAACCTTATGATTAAGGTGCTGT 1877
QY 1572 CACATCTGAGCTGCGCAATATCCAACTCAGAGCTTGGCTGCACTGCAATCAATATTTT 1631
DB 1878 CACATCTGAGCTGCGCAATATCCAACTCAGAGCTTGGCTGCACTGCAATCAATATTTT 1937
QY 1632 TCAGGACAGGATGAACCAAGCGGTTGATTAACAGCCATATGACAGATGCTCAAACTCTT 1691
DB 1938 TCAGGACAGGATGAACCAAGCGGTTGATTAACAGCCATATGACAGATGCTCAAACTCTT 1997
QY 1692 CCGATGACTGAACGAAATTTCTGTTAGACATATCTTCCGCTTGTGCTCTCATGA 1751
DB 1998 CCGATGACTGAACGAAATTTCTGTTAGACATATCTTCCGCTTGTGCTCTCATGA 2057
QY 1752 CCGGATCGCAAAAGCACTGATTTGCTGCTGCACTATGAGTCCGAGATTC 1799
DB 2058 CCGGATCGCAAAAGCACTGATTTGCTGCTGCACTATGAGTCCGAGATTC 2105

RESULT 4
US-10-140-472-358
; Sequence 358, Application US/10140472
; Publication No. US2003013888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; PRIOR APPLICATION DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550


```
RESULT 6
US-10-142-885-358
; Sequence 358, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C248
; CURRENT FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: US/10/142,885
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-885-358

Query Match      2.3%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.016;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTGAAGGGCGCGCATTTGACATGATGCTTATGTCGAACCTGATGCTGAAG 597
DB 298 HNSLSLQHVPRFWKINIKQLDELDSQNFLLAKEIGDAKFLHFLPSLIQDLDSFNFELQVYR 357
QY 598 GCTTGGCAGGCTTCCTTCCAAACATGACATGCTGCTCAACAGCAGCAGCTTTTGGAC 657
DB 358 ASNNLSQAFSSLSKLIRIGVFKELKSFNLSPLHNLQNLLEVLDTGTFIKANLSMF 417
QY 658 CAATGTTCCGATAGTGACGAGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
DB 418 KQFRKLKVIDLSVVKISPSGDSSEVGFCSNARTSVESYEPVLEQLHYFRYDKYARSCR 477
QY 718 GCGGTCATTTGGCGCTGCGATTTCCGAGCTGCTGTCGCAACGAACTGCTTCATGCTGG 777
DB 478 KKEASFMVNEVCYKYGQTLDSKNSIFPVKSSDFOHLSFLKCLNLSGMLISQTLNGSE 537
QY 778 GTAGACATGTTACAAATATATGAAGCAAGTATGCTGTGGAGCAAGCTTTGGTCACAT 837
DB 538 FQPLAEIRYLDPSNNRIDLHSTAFEBELHLEVLDISNSHYFSGEITMNLFTKXKLV 597
QY 838 GCTTTCGAGGAGCTCCTAGTGTCTGCGCAATGAGGGGCGATGCGATTTCTCTCTGCT 897
DB 598 LQKLMNDNDISSSTSTMESESLRTLEFRGNHLDVLMRSDNRXYLDLFLKMLKLELDI 657
QY 898 GCATTCGCTGTTTCTTCTTCCGAGCTTACGCGCTGCTTTCGATGAGCGCGCTTCCA 957
DB 658 SKNSLSFLPSGVDFGMPNKLKSLAKGKLSFSMKLQCLKNLETLDLSHNQLTTPER 717
QY 958 AATCCGCGACAGCTGACACTTACTTGCTTACCAAGCGTCCCAATACATGTGAAGGCC 1017
DB 718 LSNCSRLKILKILNKNQIRSLTKYFLQDAFQRLYLDLSSNKNIQMIQKTSPEPVNLNLM 777
QY 1018 GGGCAGCTGCCACGAGCTGTTCCATGCGTTTACAAAGGTTGGCGTTCGTTGAAG 1077
```

```
DB 778 LLLHNRFLLCTDAVFWVWVNVHTVETIPIPLATDVTCVBPAGAKGOSVISLDLYTCBLD 837
QY 1078 GAGCGTTTCATGAGGAGATATGCTGTGGCTTCGCCCTGTCTGCTTACTGACGCGCTTG 1137
DB 838 TNLILFSLISVSLFLMVMTASHLYFMDVWYVHFCKKIKYQRLISPDCCYDAFIYV 897
QY 1138 AATCAGACACACTTATGATGGGCTCATGACTCCTCGCAAAATTTGGCTGAACCTTTGGG 1197
DB 898 DTMDPATEVYALBELVAKLEDPREKHFNLCLERDLPQGPVLENSQISQISKTVFM 957
QY 1198 AGGAGCTCTTCTTCCGAGGATGAGAGATCTTTCTGGCACACAT 1245
DB 958 TDKYAKTENKIAFYLSHORLMDKVDVILLIFLEKFPQSKRFLQRLK 1005

RESULT 7
US-10-158-790-358
; Sequence 358, Application US/10158790
; Publication No. US20030180879A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C448
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US/10/158,790
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-158-790-358

Query Match      2.3%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.016;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTGAAGGGCGCGCATTTGACATGATGCTTATGTCGAACCTGATGCTGAAG 597
DB 298 HNSLSLQHVPRFWKINIKQLDELDSQNFLLAKEIGDAKFLHFLPSLIQDLDSFNFELQVYR 357
QY 598 GCTTGGCAGGCTTCCTTCCAAACATGACATGCTGCTTACGATGCAAGCAGCTTTTGGAC 657
DB 358 ASNNLSQAFSSLSKLIRIGVFKELKSFNLSPLHNLQNLLEVLDTGTFIKANLSMF 417
QY 658 CAATGTTCCGATAGTGACGAGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
DB 418 KQFRKLKVIDLSVVKISPSGDSSEVGFCSNARTSVESYEPVLEQLHYFRYDKYARSCR 477
QY 718 GCGGTCATTTGGCGCTGCGATTTCCGAGCTGCTGTCGCAACGAACTGCTTCATGCTGG 777
DB 478 KKEASFMVNEVCYKYGQTLDSKNSIFPVKSSDFOHLSFLKCLNLSGMLISQTLNGSE 537
QY 778 GTAGACATGTTACAAATATATGAAGCAAGTATGCTGTGGAGCAAGCTTTGGTCACAT 837
```


Db 538 FQPLAELRYLDFSNRRLDLHSTAEBELHKLVLVDLSSNSHYFQSEGITMLNFTKMLKV 597
Qy 838 GCTTTCAGGAGCCCTCTAGTGTGTCGGCCGAATGGGGCGATGCGATTCCTCGCT 897
Db 598 LQKLMNNDNDISSSTSTSESLRLEFRGNHLDVLMREGDNRVYLQFLKMLKLEBDI 657
Qy 898 GCATTCCTGCTTTTCTTCTCGAGCGCTTACGCGCTGTCTTCGATGAGCGCGTCCCA 957
Db 658 SKKSLFLPSGVFDGMPNKLKSLAKNGLKSFWSKKLQCLKNLETLDLSHNQLTVPER 717
Qy 958 AATCCCGGCGACATCGACACTTACTGCTGTACCAAGCGCTCCAATACATGTGGAAGCC 1017
Db 718 LSNCSKSLKMLILKNNQIRSLTYFLQDAPQLRYLDSLNSKIQMIOKTSFPEVNLNLM 777
Qy 1018 GGCGAGCTGCGCCAGAGCTTCCATCGGCTTTTACAGCGTGGGCGCTTCTGAAG 1077
Db 778 LLLHNRFLCTGDAVFWFVWVNHTEVTIPLYLATDVTCVPAHKGSVSLDLYTEBDL 837
Qy 1078 GAGCGTTTCATGAGGAGATATTGTGTGGCTTCCGCTGCTTACTTACGAGCGCTTG 1137
Db 838 TNLILFSLISVSLFLMVMWMTASHLYFMDVWYIYHFKAKIKGYORLLISPDCCYDAFIY 897
Qy 1138 AATCGAGCACATTAGTGGGCTCATGACTCCTGGCAAATTTGGCTGAACGCTTCGG 1197
Db 898 DTDPVATEWVLAELVAKLEDPREKHFNLCLEERDMLPGQPVLENLSQSIQLSKKTIVFW 957
Qy 1198 AGGAGTCTCTCTTCAGGAGATAGAGAGATCTTCTGGGACACAT 1245
Db 958 TDKYAKTENFKIAFYLSHORLMDKVDVILILEKPFQSKFLQLRK 1005

RESULT 8

US-10-137-871-358
; Sequence 358, Application US/10137871
; Publication No. US20030207350A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey J.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C153
; CURRENT APPLICATION NUMBER: US/10/137,871
; PRIOR FILING DATE: 2002-05-03
; PRIOR Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-871-358

Query Match 2.3%; Score 40.8; DB 12; Length 1049;

Best Local Similarity 5.4%; Pred. No. 0.016; Mismatches 473; Indels 0; Gaps 0;

Qy 538 ACCCTTGAAGGGGCGCGCATTTGACATGCTTATGTCGAACCTGATGCTGAAG 597
Db 298 HSNLSLGHVPRPWRKINKLQELDLQNLFLAKEIGDAKFLHFLPSLILQDLSPFELQYVR 357

Qy 598 GGTTCGCGAGTCTCTTCCAAATCGACTTGTCTACGACTGACGACCGTTCCTTTCAG 657
Db 358 ASMNLSQAFSSLSKSLKILIRGVFEKLSFNLSPLHNQNLQNEVLDTGNFIKIANLSMF 417
Qy 658 CAATGTTCCGATAGTGAAGCGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGAAAGTG 717
Db 418 KQFKRLKVIDLSVKNISPSGDSSEVFGCSNARTSVSEYPOVLEQHYRYDRKVASCRF 477
Qy 718 GCGGTATTGGGCTGCGACTTCCGAGCTGTGTGGTGGCAAGCAACTGCTTACGTGCG 777
Db 478 KXKEASFMSVNSSECVYGGTLLDLKSNISIFVXSDPHQLFLKCLNLISQTLNGSE 537
Qy 778 GTAGAGATGTTACATATATAGAAAGATGATCGGTTGAGGCGAAGCTTGTGCACAT 837
Db 538 FQPLAELRYLDSNNRRLDLHSTAEBELHKLVLVDLSSNSHYFQSEGITMLNFTKMLKV 597
Qy 838 GCTTTCAGGAGCCCTCTAGTGTGTCGGCCGAATGGGGCGATGCGATTCCTCGCT 897
Db 598 LQKLMNNDNDISSSTSTSESLRLEFRGNHLDVLMREGDNRVYLQFLKMLKLEBDI 657
Qy 898 GCATTCCTGCTTTTCTTCTCGAGCGCTTACGCGCTGTCTTCGATGAGCGCGTCCCA 957
Db 658 SKKSLFLPSGVFDGMPNKLKSLAKNGLKSFWSKKLQCLKNLETLDLSHNQLTVPER 717
Qy 958 AATCCCGGCGACATCGACACTTACTGCTGTACCAAGCGCTCCAATACATGTGGAAGCC 1017
Db 718 LSNCSKSLKMLILKNNQIRSLTYFLQDAPQLRYLDSLNSKIQMIOKTSFPEVNLNLM 777
Qy 1018 GGCGAGCTGCGCCAGAGCTTCCATCGGCTTTTACAGCGTGGGCGCTTCTGAAG 1077
Db 778 LLLHNRFLCTGDAVFWFVWVNHTEVTIPLYLATDVTCVPAHKGSVSLDLYTEBDL 837
Qy 1078 GAGCGTTTCATGAGGAGATATTGTGTGGCTTCCGCTGCTTACTTACGAGCGCTTG 1137
Db 838 TNLILFSLISVSLFLMVMWMTASHLYFMDVWYIYHFKAKIKGYORLLISPDCCYDAFIY 897
Qy 1138 AATCGAGCACATTAGTGGGCTCATGACTCCTGGCAAATTTGGCTGAACGCTTCGG 1197
Db 898 DTDPVATEWVLAELVAKLEDPREKHFNLCLEERDMLPGQPVLENLSQSIQLSKKTIVFW 957
Qy 1198 AGGAGTCTCTCTTCAGGAGATAGAGAGATCTTCTGGGACACAT 1245
Db 958 TDKYAKTENFKIAFYLSHORLMDKVDVILILEKPFQSKFLQLRK 1005

RESULT 9

US-10-140-805-358
; Sequence 358, Application US/10140805
; Publication No. US20030207417A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey J.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C176
; CURRENT APPLICATION NUMBER: US/10/140,805
; PRIOR FILING DATE: 2002-05-07

Db 958 TDKYAKTENKIAFYLSHORLMEKVDVILLIFLEKPFOKSKFLQLRK 1005

RESULT 11

US-10-140-923-358

; Sequence 358, Application US/10140923
; Publication No. US20030207355A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C188
; CURRENT APPLICATION NUMBER: US/10/140, 923
; PRIOR APPLICATION: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-923-358

Query Match 2.3%; Score 40.8; DB 12; Length 1049;

Best Local Similarity 5.4%; Pred. No. 0.016;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

Qy 538 ACCCTGAAGGGGCGCGCATTTGACATGATGCTTGAACCTGATGCTGAAG 597
Db 298 HNSLSQHVPPRPWKINIKLOELDSQFLAKEIGDAKFLHPLSLIQLDLSFNFELQYR 357
Qy 598 GGTTCGGCAGGTTCTTCCAAACATGACCTTGTCTTACAGCTGACAGCCGTTTGTGAC 657
Db 358 ASNMLSQAFSLKSLKILIRGVYFKELKSFNLPLNLQNLVLDLGNFIKIANLSMF 417
Qy 658 CAATGTTCCGATAGTGAAGGATGCGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
Db 418 KQFRLKVLIDLSVSKISPSGDSSEVFCNSARTSVSEYPOVLEQLHYFYDYKARSQRP 477
Qy 718 GCGGTCAATGGCGTGCATTTCCGACACTGCTGTGTCGAACCAAGACTGATGCTGAGG 777
Db 478 KKEASMSVNSCYVGGTDLDSKNSIFVYKSSDPQHLFLKLNLSGULISQTLNGSB 537
Qy 778 GTAGACAGATGTTCAATATATGAAGCAAGTATGCTGTGAGGCAAGCTTGTGTCACAT 837
Db 538 FQPLAELRYIDFSNNRDLHSTAFBELHLEVLDISNSHYPGSEGITMNLNFTKLV 597
Qy 838 GCTTTCAGGACGCTCTAGTGTGTCGCGGAATGGGGCGATGATTTCTCTGCT 897
Db 598 LQKLMDNDISSSTRHSESLRTEFRGNHLDVLRBGDNRYIQLFKNLLKLELDI 657
Qy 898 GCATTCGCTGTTTCTTCTCGACGCTTACGCGCTGTCTTCGATGAGAGCCGTTCCA 957
Db 658 SKNSLSLBPBGVDPMPNLKNSLAKNGLSKSWKLCQLKLELITLIDLSHNOULTIVPER 717
Qy 958 AATCCCGGACAGTCACTTACTGCTTACCAAGCGCTCCAAATCATATGATGGAAGCC 1017
Db 718 LSNCSRLKLVLLKNNQIRSLTYFLQDAPQLRLYLDLSNKNQIQKTSPPENVANLNLKM 777

Qy 1018 GGCAGCTGCACCAAGACTGTTCCATGCGCTTACAGGCTTGGCGCTTCTGAAG 1077
Db 778 LLHNRFLCTCAVAFVWVWVHTETIPLATIDVTCVGGAKKGSVSLDYTEL 837
Qy 1078 GACGTTTTCATGAGCAATATGTTGTTGCTTCCCTGCTTCTTACTTACAGGCTTG 1137
Db 838 TNLIFSLISVSLFLMWMVMTASHLYFWDVWVYIYHCKAKIKGYRLISPDCCVDAPIV 897
Qy 1138 AATCAGACACTTATGTTGGCTCATGACTCTTGGCAATTTGGCTGAACGTTTCCGG 1197
Db 898 DTDPAVTEVWVLELAKLEDPREKHNLCLEBDMLPGQPVLENLSQSISQLSKTVFWM 957
Qy 1198 AGGAGTCTTCTTCTCGAGGATGAGAGATCTTCTGCGACACAT 1245
Db 958 TDKYAKTENKIAFYLSHORLMEKVDVILLIFLEKPFOKSKFLQLRK 1005

RESULT 12

US-10-141-756-358

; Sequence 358, Application US/10141756
; Publication No. US20030207359A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C200
; CURRENT APPLICATION NUMBER: US/10/141, 756
; PRIOR APPLICATION: 2002-05-08
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-756-358

Query Match 2.3%; Score 40.8; DB 12; Length 1049;

Best Local Similarity 5.4%; Pred. No. 0.016;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

Qy 538 ACCCTGAAGGGGCGCGCATTTGACATGATGCTTGAACCTGATGCTGAAG 597
Db 298 HNSLSQHVPPRPWKINIKLOELDSQFLAKEIGDAKFLHPLSLIQLDLSFNFELQYR 357
Qy 598 GGTTCGGCAGGTTCTTCCAAACATGACCTTGTCTTACAGCTGACAGCCGTTTGTGAC 657
Db 358 ASNMLSQAFSLKSLKILIRGVYFKELKSFNLPLNLQNLVLDLGNFIKIANLSMF 417
Qy 658 CAATGTTCCGATAGTGAAGGATGCGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
Db 418 KQFRLKVLIDLSVSKISPSGDSSEVFCNSARTSVSEYPOVLEQLHYFYDYKARSQRP 477
Qy 718 GCGGTCAATGGCGTGCATTTCCGACACTGCTGTGTCGAACCAAGACTGATGCTGAGG 777
Db 478 KKEASMSVNSCYVGGTDLDSKNSIFVYKSSDPQHLFLKLNLSGULISQTLNGSB 537

QY 778 GTAGACGATGTTACAAATATATGAAGCAAGTAGTCGTTGGAGGCAACTTTGGCACAT 837
Dd 538 FQPLAELRYVDFSNRRDLHSHTAPEELHKEVLDISNSHYPQSEBGIITHINLFPKMLXV 597
QY 838 GCTTTACGAGCAGCTCTCCTAGTCGTGGCCGAATGGGGCGAGCGCATTCCTCTCTCT 897
Dd 598 LQKLMDNDNDISSSTRTMSEBSRLTFEPFGHLDVLRLREGDNRYLQFLKMLKLELDI 657
QY 898 GCATTCTGCTGTTTTCTTCTCTCGAGGATACGGCCGTCGTGCATAGAGCGCCCTTCCA 957
Dd 658 SKNSLSFLPSGVFQDMPPNLYKLSLAKKGLSFSKKIQLCLNLETTLDLSHQLTIVBER 717
QY 958 AATCCCGGACAGTCGACACTTACTTGCTGTACCAAGCGCTGCAATACATGTGAAAGCC 1017
Dd 718 LSNCSRSLKMLILKKNQJRLSTKYPQLDQAFQRLYDLSSNKIKOMIQTSPENVALNLM 777
QY 1018 GGGCAGCTGCCACCGAAGCTGTTCCATCGCCGTTTCAACAGCTGGCCGCGCTTTGAAG 1077
Dd 778 LLLHHNRFELCTCDAWFVWVMMHNEVTLRYLATDVTGCPGAKHKSQVSIIDLTYLCELD 837
QY 1078 GACGGTTTTCAGACGAGATATTGTTGCTGCTGCTGCTGCATTAATCTAGAGCCCTTG 1137
Dd 838 TNNLLFSLSISVSLFMMVMTASHIYFPDWVIYHFCAKIKGYQRLLSPDCDYAFIV 897
QY 1138 AAATCGAGACATTAGTGGGCTCATGACTCTCCGCAATTTGGCTGAACGTTGGCG 1197
Dd 898 DTKDPAYEWVLAELVAKLDEPREKHFNILCLEEROMLPGQPVLENLSOSIOLSKTVEVM 957
QY 1198 AGGAGTCTCTCTTCAGGAGTATAGAGAGATCTTTGCGGACACAT 1245
Dd 958 TDKRAKTENFKIAFLYLSHORLMDKENVADYIILFLEKRPQKSFQDLRK 1005

```

RESULT 13
US-10-141-759-358
; Sequence 358, Application US/10141759
; Publication No. US20030207361A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C197
; CURRENT APPLICATION NUMBER: US/10/141,759
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-141-759-358

```

Query Match 2.3%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.016;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

Db	298	HSNSLQHVPRWFNINKLQELBDSQNFALKEIGDAFHFHFLPSLJQLDLSFNFELQVFR	357
OY	598	GGTTGGGAGGCTTCCTTCCAAACATGACTGTGCTACGACCTGCAGACCGCTTTTGGAC	657
Db	358	ASNMLSQAFSLSKLKLIRIGVYFKELKSFNLSPLNLONLEVLDTGTFIKIANLSMF	417
OY	658	CAATGTTCCGATAGTAGGAGGATGGCTTCTTCCGAGGATGTTCTTAAGCCAAAGTG	717
Db	418	KQFKRLKVIDLSVWKISPSBDSBVGCNSAARTSVSEYEQVLEQLHFYDKYARBCRF	477
OY	718	GCGETCATTTGGCGCTGCATTTCCGACCTGCTGGGCAAAAGAACTGTTATGCTGGG	777
Db	478	KNKASFMVSVBESCYKQGTLDSIKNSIFPVKSDPGLHFLKLYLNSGMLISQTLNGSE	537
OY	778	GTAAGCATGTTACATATATATGAAGCAAGATGCTGTTGGAGCAAGCTTGTGCAT	837
Db	538	FQPLAEIRYUDEFNSNNRDLHSTHFAEBYHKLVLJDLSNHHYFQSEGITMHLNFTKLV	597
OY	838	GCTTCAAGGAGCGCTCCAGTGCTGAGCGGAAATGGGGGAGATGCATTTCCCTGCT	897
Db	598	LQKLMMDNDLSSSTSHMESSESRTEBFKNHLDVIMRBEDNRYLDLFLKLYLLELDI	657
OY	898	GCATTGCTGTTGTTTTCTTCCGAGCGTTACGCGCTGCTTCCGATGAGCGCTTCCA	957
Db	658	SKNSLFLPSGVFQMPFNILKNLSLANKGLKSFWKLYQCLKNLETLDSLHQTLTVBER	717
OY	958	AATCCCGGACAGTGCACATTTAGTGTCTACCAAGCGCTCCATATCATGTGGAAGCC	1017
Db	718	LSNCSRSILKNILKNNOIRSLTKFLDAPOLRYLDSLNSKIOIMIOKTSPPENVLNLKM	777
OY	1018	GGGAGCGTCCAGCGAGCTGTTCCATCGGCTTACAACGGTTGGCGGCTTGTGAAG	1077
Db	778	LLLHHNRLCTCDVWVFWVWVWVWVWVHLEVTIIPYATDVTCVGRANHGQSVISLDLYTCEL	837
OY	1078	GACGGTTTTCATGACGAGATATGTTGTTGCTTCCGCTGCTGCTATTTACTCAGCGCTG	1137
Db	838	TNLLFLSLISIVSLFWMMWMTASHLYPMDWUYIHFCAKIKYQVRLISPDCCVAFIVY	897
OY	1138	AAATCGAGACATTTAGTGGGCTCATGACTCTCCGCAAAATTTGGCTGMAACGGTTGGG	1197
Db	898	DTKQPAVTEWVLAELVAKLEDPREKHFNLCLSEERDMLPGQVLENLSQS10LSKKTIVM	957
OY	1198	AGGAGTCTCTCTCAGGAGTAGAGGATCTTTTGGGACACAT	1245
Db	958	TDKAKTENFKIAFYLSHORLMDKVDVYIILFEKRPQSKFQDLK	1005

```

RESULT 14
US-10-123-155-358
; Sequence 358, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30

```


OY 1198 AGGAGTCTTCTTCAGGATAGAGGATCTTCTGGGACACAT 1245
Db 958 TDKYAKTENFKIAFYLSHQRLMDKVDVILIFLEKPFQSKFQLQLRK 1005

Search completed: November 23, 2003, 21:22:48
Job time : 560.252 secs

PT untranslatable nucleic acid homologous to a gall disease-causing gene
PS Claim 2, Page 45; 49pp; English.

XX The patent discloses a method for producing plants resistant to gall
CC disease, induced by Agrobacterium infection, by transformation with a
CC bacterial resistance (BR) construct capable of eliciting co-suppression.
CC The construct encodes an untranslatable RNA molecule, that is highly
CC homologous to the tumor or gall disease-causing genes *iaam*, *iaah* and
CC *ipt* of *Agrobacterium tumefaciens*. These oncogenes are modified by
CC introduction of premature termination codons or frameshift mutations, to
CC inhibit their expression. This method is used to reduce susceptibility
CC of plants to gall diseases. The present DNA sequence is the *Agrobacterium*
CC tumefaciens T-DNA (transferred DNA) gene *iaam*, required for auxin
CC production. It encodes an enzyme that converts tryptophan into indole
CC acetamide, essential for auxin synthesis. The overproduction of these
CC plant growth hormones results in gall formation.

XX Sequence 2268 BP; 543 A; 520 C; 606 G; 599 T; 0 other;

Query Match 100.0%; Score 2268; DB 21; Length 2268;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGTCACCTTCACTCTCTGATTAACCACTGCGATCATCTCCCAACCAATGCTGAT 60
DB 1 ATGTCACCTTCACTCTCTGATTAACCACTGCGATCATCTCCCAACCAATGCTGAT 60
QY 61 CTGACATATGTCATTAAGCGGATGATGGAACCGAGGCTTCGATGCTTCTAGAA 120
DB 61 CTGACATATGTCATTAAGCGGATGATGGAACCGAGGCTTCGATGCTTCTAGAA 120
QY 121 CGAAGACCTTCTAGGGAAGAGGATTAATCAAGCTTCAACCGAGTCAACGCTGGGTTA 180
DB 121 CGAAGACCTTCTAGGGAAGAGGATTAATCAAGCTTCAACCGAGTCAACGCTGGGTTA 180
QY 181 GCTTGAAGAGGCGCGGATGCTGCTCCGAGATCTCAGTGTGGAAGATGCA 240
DB 181 GCTTGAAGAGGCGCGGATGCTGCTCCGAGATCTCAGTGTGGAAGATGCA 240
QY 241 GTTCTCTCCGCTTATATCTATTTGCAAGAAATTCCTGGGCGGATCTTGAATCGAA 300
DB 241 GTTCTCTCCGCTTATATCTATTTGCAAGAAATTCCTGGGCGGATCTTGAATCGAA 300
QY 301 CTTTGGGCGGCGGACAGTGAAGTGTCTCTGTCATCACTTGAACCAATTTGCATG 360
DB 301 CTTTGGGCGGCGGACAGTGAAGTGTCTCTGTCATCACTTGAACCAATTTGCATG 360
QY 361 GATTTCCTCGAAGACATTAATTCAGAGCCCTGTTTTGCTGAAGCGGTTAAAGATGCA 420
DB 361 GATTTCCTCGAAGACATTAATTCAGAGCCCTGTTTTGCTGAAGCGGTTAAAGATGCA 420
QY 421 CCGATTGATCTTATGATCTTCTGAGCCATTTCAATCTCTAAGATGCGGCTTTCGAAC 480
DB 421 CCGATTGATCTTATGATCTTCTGAGCCATTTCAATCTCTAAGATGCGGCTTTCGAAC 480
QY 481 CTGCGCAATGCGCTGTATGAGAAATGAGATGAATGCGTTTACCGGTTTACCAATACC 540
DB 481 CTGCGCAATGCGCTGTATGAGAAATGAGATGAATGCGTTTACCGGTTTACCAATACC 540
QY 541 CTTGAAGAGGCGGCGGATTTGACATGCTTAAGTGTGGAACCTGATGCTGAAGGT 600
DB 541 CTTGAAGAGGCGGCGGATTTGACATGCTTAAGTGTGGAACCTGATGCTGAAGGT 600
QY 601 TCGGCAAGTCTCTTCCAAATGCACTTCTCTGACATGCAAGCCGTTTTCGACCA 660
DB 601 TCGGCAAGTCTCTTCCAAATGCACTTCTCTGACATGCAAGCCGTTTTCGACCA 660
QY 661 TGTTCGATATGAGCGGATCGGCTTCTTCCGAGGATGTTCTTAAGCCGAAAGTGGC 720
DB 661 TGTTCGATATGAGCGGATCGGCTTCTTCCGAGGATGTTCTTAAGCCGAAAGTGGC 720
```

```
QY 721 GTCATTGGCGCTGGCATTTCCGAGCTGATGAGCAACCAACGCTTCACTGCTGGGTA 780
DB 721 GTCATTGGCGCTGGCATTTCCGAGCTGATGAGCAACCAACGCTTCACTGCTGGGTA 780
QY 781 GACGATGTTAACAATATATGAGCAAGATGATCGTGTGAGGCAAGCTTTGGTCAATGCT 840
DB 781 GACGATGTTAACAATATATGAGCAAGATGATCGTGTGAGGCAAGCTTTGGTCAATGCT 840
QY 841 TTGAGGAGCGCTCTTATGCTGCTGCGGCAATGAGGCGCATGCGATTTCTCTGCTGCA 900
DB 841 TTGAGGAGCGCTCTTATGCTGCTGCGGCAATGAGGCGCATGCGATTTCTCTGCTGCA 900
QY 901 TTCTGCTGTTTTTCTCTCTCGAGCGCTTACGCGCTGCTCTGATGAGCGCTTCCCAAT 960
DB 901 TTCTGCTGTTTTTCTCTCTCGAGCGCTTACGCGCTGCTCTGATGAGCGCTTCCCAAT 960
QY 961 CCGGACACAGTCGACATTAATTTGCTTACCAAGGCGTCCAAATCATGAGAAAGCCGG 1020
DB 961 CCGGACACAGTCGACATTAATTTGCTTACCAAGGCGTCCAAATCATGAGAAAGCCGG 1020
QY 1021 CAGCTGCAACCGAAGCTTTCATCGCGTTTACAGGTTGGCGCTTCTTGAAGGAC 1080
DB 1021 CAGCTGCAACCGAAGCTTTCATCGCGTTTACAGGTTGGCGCTTCTTGAAGGAC 1080
QY 1081 GATTTTCAGAGGAGATATGCTTGGCTTGCCTGTGCTATTACTCAGGCTTGA 1140
DB 1081 GATTTTCAGAGGAGATATGCTTGGCTTGCCTGTGCTATTACTCAGGCTTGA 1140
QY 1141 TCAGACACATTAAGTGGGCTCATGACTCTGGCAAAATTTGGCTGAACCGTTTCCGAGG 1200
DB 1141 TCAGACACATTAAGTGGGCTCATGACTCTGGCAAAATTTGGCTGAACCGTTTCCGAGG 1200
QY 1201 GAGTCTCTCTCTCAGGATAGAGAGATCTTCTGGGCAACATCTCTGCTGGTAA 1260
DB 1201 GAGTCTCTCTCTCAGGATAGAGAGATCTTCTGGGCAACATCTCTGCTGGTAA 1260
QY 1261 CATGAGATTTTCTCATGATTTGGAGACCTATTCAGCTAATGGAATGGAATCTGGCGGG 1320
DB 1261 CATGAGATTTTCTCATGATTTGGAGACCTATTCAGCTAATGGAATGGAATCTGGCGGG 1320
QY 1321 TTTGGTCAAGTTTTTGAAGCGGTTTATGAGATCTCTGGTTCATCAACGATAT 1380
DB 1321 TTTGGTCAAGTTTTTGAAGCGGTTTATGAGATCTCTGGTTCATCAACGATAT 1380
QY 1381 GAAGAAATCAAGGAGATGTCCTGGAAGAAATCTCAAGACTTCAGTGGATCGATCT 1440
DB 1381 GAAGAAATCAAGGAGATGTCCTGGAAGAAATCTCAAGACTTCAGTGGATCGATCT 1440
QY 1441 GAAGGATTAAGGATGCTGTGAGCAGCGCATATGCGATGTTCAAGTCAGGCGATT 1500
DB 1441 GAAGGATTAAGGATGCTGTGAGCAGCGCATATGCGATGTTCAAGTCAGGCGATT 1500
QY 1501 CAGAGAAAGACCAAAATTAAGATTAAGCTTAAGCGGATCTGAACCTTTATGAT 1560
DB 1501 CAGAGAAAGACCAAAATTAAGATTAAGCTTAAGCGGATCTGAACCTTTATGAT 1560
QY 1561 AAGTGTGTGTCATCTGATGCTGCAATATCAACTCAGGATGCTGCAATGCGAT 1620
DB 1561 AAGTGTGTGTCATCTGATGCTGCAATATCAACTCAGGATGCTGCAATGCGAT 1620
QY 1621 ACCAATATTTTGAAGCAGAGTGAACCAAGCGGTTGATTAACAGCATATGACAGATCG 1680
DB 1621 ACCAATATTTTGAAGCAGAGTGAACCAAGCGGTTGATTAACAGCATATGACAGATCG 1680
QY 1681 TCAGAACTCTTCTCTGATGATGACAGAAATTTGCTGATGACATATCTCCGCTCTTGT 1740
DB 1681 TCAGAACTCTTCTCTGATGATGACAGAAATTTGCTGATGACATATCTCCGCTCTTGT 1740
QY 1741 GTCTCATGAGCGGATTCGAAAGAGATGATGCTGATGATGATGATGATGATGATGAT 1800
DB 1741 GTCTCATGAGCGGATTCGAAAGAGATGATGCTGATGATGATGATGATGATGATGAT 1800
QY 1801 AATGTAAGGATTAATGCTCATCACTTAATCACTGGAAGAGACATCCCAAGCTGTTG 1860
```

DB 1801 AATGTAAGAGTGTAGCTCATCAGTTATACATGGAGAGACATCCCAAGCTGTG 1860
QY 1861 GCGGTCCCGCAAAAAAGAGCATTTATGCTGTGGGGAGCCAAATTTGAGATCTTC 1920
DB 1861 GCGGTCCCGCAAAAAAGAGCATTTATGCTGTGGGGAGCCAAATTTGAGATCTTC 1920
QY 1921 CCGGCGTTTGCAGCAGCATTTATCTGCTGCGCTGATACGACCAAAATTTATTC 1980
DB 1921 CCGGCGTTTGCAGCAGCATTTATCTGCTGCGCTGATACGACCAAAATTTATTC 1980
QY 1981 CATGATTTGCTTACAGAGCAATATGCGGGGAGCTTTCAACTCAACCGGCGTGTAG 2040
DB 1981 CATGATTTGCTTACAGAGCAATATGCGGGGAGCTTTCAACTCAACCGGCGTGTAG 2040
QY 2041 GATTTTATTTCTAAGAACTTTCTTTCAGACACTGACACCGCTATATATACCGAGTT 2100
DB 2041 GATTTTATTTCTAAGAACTTTCTTTCAGACACTGACACCGCTATATATACCGAGTT 2100
QY 2101 TACTTGGCGGTTGCAAGTTGCTTCTTCAAGAGTGAATGGGAGGCTATTCAGACC 2160
DB 2101 TACTTGGCGGTTGCAAGTTGCTTCTTCAAGAGTGAATGGGAGGCTATTCAGACC 2160
QY 2161 GCGGTGAACGCGCTGTGTGCAATTTATCCAAATTTGTGAGGCAATTTTGCAAGGCGCAT 2220
DB 2161 GCGGTGAACGCGCTGTGTGCAATTTATCCAAATTTGTGAGGCAATTTTGCAAGGCGCAT 2220
QY 2221 CCTCTGCAACACTCTTGAAGAGATATACCTCCGCTGAAATTTG 2268
DB 2221 CCTCTGCAACACTCTTGAAGAGATATACCTCCGCTGAAATTTG 2268

RESULT 2

ID AAN50182 standard; DNA; 24596 BP.

AAN50182;

XX 25-MAR-2003 (updated)
DT 17-OCT-1991 (first entry)XX Complete nucleotide sequence of the T-DNA region of the octopine Ti
DE plasmid pTi15955.XX Plant vector; transformation-inducing principle (TiP) gene;
KM octopine Ti plasmid; ss.

XX Agrobacterium tumefaciens ATCC 15955.

XX EPI45338-A.

XX 19-JUN-1985.

XX 16-NOV-1984; 84BP-0307969.
XX 18-NOV-1983; 83US-0553786.XX (AGRK) AGRIGENETICS RES ASSOC LTD.
XX (LUBR) LUBRIZOL GENETICS INC.XX Barker RF, Kemp JD;
XX WPI, 1985-148223/25.XX New DNA vectors contg. T-DNA sequence of octopine Ti plasmid -
PT for expression in plant cells to confer desirable properties to
XX plants and their cells

XX Claim 28, Fig 1; 87p; English.

XX The inventors claim a vector contg. a transformation-inducing
CC principle (TiP) gene from Ti plasmid pTi15955. The sequence of the
CC T-DNA of the octopine-type Ti plasmid has fourteen open reading

CC frames bounded by eukaryotic promoters, ribosome binding sites, and
CC polyadenylation sites. With the vectors, expression of structural
CC foreign genes in plant cells is promoted. The gene esp. encodes an
CC insecticidal toxin identical to or derived from the crystal protein
CC of *Bacillus thuringiensis*.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX

SQ Sequence 24596 BP; 6534 A; 5510 C; 5793 G; 6759 T; 0 other;

Query Match 98.4%; Score 2230.6; DB 6; Length 24596;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 2258; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 1 ATGTCACTTCACTCTCTCTGATTAACAGTGGATCATCTCCCAACCAAAATGTGTAT 60
DB 5809 ATGTCACTTCACTCTCTCTGATTAACAGTGGATCATCTCCCAACCAAAATGTGTAT 5868
QY 61 CTGACAAATGTCGATTAAGCGGATGAATGACCCGAGGCTTCCGATCCTTTAGAA 120
DB 5869 CTGACAAATGTCGATTAAGCGGATGAATGACCCGAGGCTTCCGATCCTTTAGAA 5928
QY 121 CGAAGAGCTTCTAAGGGAAGAGATTAATCAAGCTCCACCGAGTGCAGCGTGGGTTA 180
DB 5929 CGAAGAGCTTCTAAGGGAAGAGATTAATCAAGCTCCACCGAGTGCAGCGTGGGTTA 5988
QY 181 GCTTGCAAAAGGCTGGCGATGTGCTCCCGAGATCTCAGCTGTGGAAAGTATGA 240
DB 5989 GCTTGCAAAAGGCTGGCGATGTGCTCCCGAGATCTCAGCTGTGGAAAGTATGA 6048
QY 241 GTTCTCTCCGCTTATATCTATATTGGCAAAATTTCTGGGGGAGTACTTGAATGAAA 300
DB 6049 GTTCTCTCCGCTTATATCTATATTGGCAAAATTTCTGGGGGAGTACTTGAATGAAA 6108
QY 301 CCTTGGGCGCGGCAACAGTGAAGTGTCTGCTTCCATGACTTGACACCATTTGCATG 360
DB 6109 CCTTGGGCGCGGCAACAGTGAAGTGTCTGCTTCCATGACTTGACACCATTTGCATG 6168
QY 361 GATTTCTCGAAGACAACTAATCCAAAGCCCTTTTCTGTGCGGTTAAAGATGTGA 420
DB 6169 GATTTCTCGAAGACAACTAATCCAAAGCCCTTTTCTGTGCGGTTAAAGATGTGA 6228
QY 421 CCGATTGATCTTATGTCATTTGCGCATTTTCAATCTTAAGACTGCGGCTTTCGAAC 480
DB 6229 CCGATTGATCTTATGTCATTTGCGCATTTTCAATCTTAAGACTGCGGCTTTCGAAC 6288
QY 481 CTGCCAATGCCGCTGTACAGAAATGSCACAGTAAATGCGTTACCGGTTTACATTAAC 540
DB 6289 CTGCCAATGCCGCTGTACAGAAATGSCACAGTAAATGCGTTACCGGTTTACATTAAC 6348
QY 541 CTTGAAGGGGCGGTGCAATTTGACATGATGCTTAATGTGCAAACTGANTCGGAAGGCT 600
DB 6349 CTTGAAGGGGCGGTGCAATTTGACATGATGCTTAATGTGCAAACTGANTCGGAAGGCT 6408
QY 601 TCGGCAAGTCTTCTTCCAACTGCTCTTCAAGTGTGAGAGCCGTTTTTGAACAA 660
DB 6409 TCGGCAAGTCTTCTTCCAACTGCTCTTCAAGTGTGAGAGCCGTTTTTGAACAA 6468
QY 661 TGTTCGATAGTGAAGCATGCGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 720
DB 6469 TGTTCGATAGTGAAGCATGCGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 6528
QY 721 GTCATTTGGGCTGGCAATTTCCGACTCGTGTGAGCAAGCACTGCTTATAGTGGGGTA 780
DB 6529 GTCATTTGGGCTGGCAATTTCCGACTCGTGTGAGCAAGCACTGCTTATAGTGGGGTA 6588
QY 781 GAGCATGTATACATATATGAAGCAAGTGAATGCTGTGAGGCAAGCTTTGTCAATGCT 840
DB 6589 GAGCATGTATACATATATGAAGCAAGTGAATGCTGTGAGGCAAGCTTTGTCAATGCT 6648
QY 841 TTCAGGACGCTCTTATGTCGTGCGCAATGAGGCGCATGCGATTTCTCTCTGTGCA 900
DB 6649 TTCAGGACGCTCTTATGTCGTGCGCAATGAGGCGCATGCGATTTCTCTCTGTGCA 6708

QY 901 TTGCGCTGTTTTCTTCTCTGAGCGCTTAAAGCGCTGTCTGATGAGCGCGTTCCCAAT 960
 DB 6709 TTTGCTGTTTTTTCTTCTCTGAGCGCTTAAAGCGCTGTCTGATGAGCGCGTTCCCAAT 6768
 QY 961 CCGGCGACAGTCGACACTTACTTGTGTCTACCAAGCGCTTCAATATCATGTGAAAGCGGG 1020
 DB 6769 CCGGCGACAGTCGACACTTACTTGTGTCTACCAAGCGCTTCAATATCATGTGAAAGCGGG 6828
 QY 1021 CAGTCCGACCGAAGCTTTCATCGCGTTTACAAGCGTTGGCGCTTCTTGAAGAC 1080
 DB 6829 CAGTCCGACCGAAGCTTTCATCGCGTTTACAAGCGTTGGCGCTTCTTGAAGAC 6888
 QY 1081 GGTTCATAGAGAGATTTGTTGGCTTGGCTGCTGCTGCTTACTTCAAGCGCTTGA 1140
 DB 6889 GGTTCATAGAGAGATTTGTTGGCTTGGCTGCTGCTGCTTACTTCAAGCGCTTGA 6948
 QY 1141 TCAGACACATTAAGTGGGCTCATGACTCCGCAAAATTTGGCTGAACCGTTTCGGAGG 1200
 DB 6949 TCAGACACATTAAGTGGGCTCATGACTCCGCAAAATTTGGCTGAACCGTTTCGGAGG 7008
 QY 1201 GAGTCTTCTCTCAAGGATAGAGAGATCTTTCTGGGCAACATCTCTGCTGTGT-AA 1259
 DB 7009 GAGTCTTCTCTCAAGGATAGAGAGATCTTTCTGGGCAACATCTCTGCTGTGTAA 7068
 QY 1260 ACATGAGTTTCTCTCAATGAGGACCTTATTCAGCTAATGGGAATAGGATCTGGGGG 1319
 DB 7069 ACATGAGTTTCTCTCAATGAGGACCTTATTCAGCTAATGGGAATAGGATCTGGGGG-GG 7127
 QY 1320 GTTTGCTCAGTTTTTGAAGCGGTTTATTAAGATCTCTCGCTGTGCTATCAACGAGTA 1379
 DB 7128 GTTTGCTCAGTTTTTGAAGCGGTTTATTAAGATCTCTCGCTGTGCTATCAACGAGTA 7187
 QY 1380 TGAAGAAATCAGCGGATGTGCTTGAAGATTCAGAACTTCAAGCTGAGATCGATC 1439
 DB 7188 TGAAGAAATCAGCGGATGTGCTTGAAGATTCAGAACTTCAAGCTGAGATCGATC 7247
 QY 1440 TGAAGTGTAAAGGCTGTCTGTGAGCAAGCGCATATGCGATTTCAAGTCAAGGGGAT 1499
 DB 7248 TGAAGTGTAAAGGCTGTGTGTGTGAGCAAGCGCATATGCGATTTCAAGTCAAGGGGAT 7307
 QY 1500 TCAGAGGAAAGACAAATAAAGATTAAGGCTTAAAGCGGATATCTGAACCTTTATGA 1559
 DB 7308 TCAGAGGAAAGACAAATAAAGATTAAGGCTTAAAGCGGATATCTGAACCTTTATGA 7367
 QY 1560 TAAAGTGTGTGCTCATCTGAGCTCGCAATATCAACTCAGGATTTGCTGATCGCA 1619
 DB 7368 TAAAGTGTGTGCTCATCTGAGCTCGCAATATCAACTCAGGATTTGCTGATCGCA 7427
 QY 1620 TACCAATATTTTCAAGGACCAAGTGAACCAAGCGGCTTAAAGCGCATATGACGAGATC 1679
 DB 7428 TACCAATATTTTCAAGGACCAAGTGAACCAAGCGGCTTAAAGCGCATATGACGAGATC 7487
 QY 1680 GTCAAACTCTTCTGTGATGACTGAACGAAATTTCTGTTAGACCATATCTCCCGTCTTG 1739
 DB 7488 GTCAAACTCTTCTGTGATGACTGAACGAAATTTCTGTTAGACCATATCTCCCGTCTTG 7547
 QY 1740 TGTCTCATGAGACGGGATCCGCAAAAGAGTGTATGCTGCTGATCATGATGCGAGATCC 1799
 DB 7548 TGTCTCATGAGACGGGATCCGCAAAAGAGTGTATGCTGCTGATCATGATGCGAGATCC 7607
 QY 1800 GAATGTAAAGTGTATGTGCTCATGATTAACATGAGAGAGACCTCCCAAGAGCTT 1859
 DB 7608 GAATGTAAAGTGTATGTGCTCATGATTAACATGAGAGAGACCTCCCAAGAGCTT 7667
 QY 1860 GGGCGTCCCGACAAAAAAGAGCATATGTCTGCTGCGGACGCAATTTGAGATCTTT 1919
 DB 7668 GGGCGTCCCGACAAAAAAGAGCATATGTCTGCTGCGGACGCAATTTGAGATCTTT 7727
 QY 1920 CCGGCGCTTTGGCCAGACCTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATTC 1979
 DB 7728 CCGGCGCTTTGGCCAGACCTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATTC 7787
 QY 1980 ACATGATTTGGCTTACACAGAGATGCGGGGAGCTTTCAAACTCAACGGCGCTGTGA 2039

DB 7788 ACATGATTTGGCTTACAGACAGAGATCCGGGGAGCTTTCAAACTCAACGGCGTGTGA 7847
 QY 2040 GATTTTTTATTTGAGAACTTTCTTTCAAGCAGTCGACACGGCTAATGATACCGGAGT 2099
 DB 7848 GATTTTTTATTTGAGAACTTTCTTTCAAGCAGTCGACACGGCTAATGATACCGGAGT 7907
 QY 2100 TTACTTTGGGGGTGAGTGTGTTCTTCAAGTGAATGGGTGAGAGGCTTATTCAGAC 2159
 DB 7908 TTACTTTGGGGGTGAGTGTGTTCTTCAAGTGAATGGGTGAGAGGCTTATTCAGAC 7967
 QY 2160 CCGGTGAACGCGCTGTGCAATTAATCCAAATTTGAGAGCGATTTTGGCAAGGCA 2219
 DB 7968 CCGGTGAACGCGCTGTGCAATTAATCCAAATTTGAGAGCGATTTTGGCAAGGCA 8027
 QY 2220 TCCTTCGAACACTCTTGAAGAGATATACTACCGCAGTAAGAAATTAG 2268
 DB 8028 TCCTTCGAACACTCTTGAAGAGATATACTACCGCAGTAAGAAATTAG 8076

RESULT 3
 AAN50226
 ID AAN50226 standard; DNA; 24593 BP.
 XX
 AC AAN50226;
 XX
 DT 25-MAR-2003 (updated)
 DT 16-OCT-1991 (first entry)
 XX
 DE Sequence of opine synthase gene.
 XX
 KM Plant vector; Ti plasmid; T-DNA; se.
 XX
 OS Agrobacterium tumefaciens Ti plasmid.
 XX
 PN EP140556-A.
 XX
 PD 08-MAY-1985.
 XX
 PF 12-SEP-1984; 84EP-0306233.
 XX
 PR 14-SEP-1983; 83US-0532280.
 XX
 PA (AGRK) AGRIGENETICS RES ASSOC LTD.
 PA (LUBR) LUBRIZOL GENETICS INC.
 XX
 PI Dahl GA, Sutton DW, Barker RF;
 XX
 DR WPI; 1985-112088/19.
 XX
 PT plasmid contg. opine synthase gene for selection - and foreign
 PT DNA, useful as vector for transforming plant cells
 XX
 PS Disclosure; Page 212-217; 69pp; Japanese.
 XX
 CC Plant cells (and protoplasts) and plasmid contg. the DNA fragment
 CC which includes an opine synthase gene plus a gene for antibiotic
 CC resistance are claimed. These plasmids provide max. efficiency for
 CC transfer of foreign genes and can be amplified in the plant genome.
 CC They do not contain genes specifying tumour formation and will not
 CC spread antibiotic resistance throughout the plant population.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 XX
 SQ Sequence 24593 BP; 6554 A; 5493 C; 5787 G; 6759 T; 0 other;
 XX

Query Match 98.1%; Score 2225.8; DB 6; Length 24593;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 2255; Conservative 0; Mismatches 12; Indels 2; Gaps 2;
 QY 1 ATGTGAGCTTCACTCTCTTGAATAACAGTGGATCATCTCCCAACCAAAATGTGGAT 60
 DB 5809 ATGTGAGCTTCACTCTCTTGAATAACAGTGGATCATCTCCCAACCAAAATGTGGAT 5868

QY 61 CTGCAATGTCGATTAAGCGCGATGAATTGGAACCGCGAGGGTTTCCGATGCCCTTTAGAA 120
DB 5869 CTGCAATGTCGATTAAGCGCGATGAATTGGAACCGCGAGGGTTTCCGATGCCCTTTAGAA 5928
QY 121 CGAAGACCTTCTAGGGGAGGAGATTACTCAAAATCTCCAGATGCGAGCTGAGGTTA 180
DB 5929 CGAAGACCTTCTAGGGGAGGAGATTACTCAAAATCTCCAGATGCGAGCTGAGGTTA 5988
QY 181 GCTTGGAAAAAGCTGGCGGATGTCCTTCCCGAGATCTCAGCTGGTGAAGTNGCA 240
DB 5989 GCTTGGAAAAAGCTGGCGGATGTCCTTCCCGAGATCTCAGCTGGTGAAGTNGCA 6048
QY 241 GTTCTCTCCGCTTATCTATATATGCGAAGAAATTCCTGGGGCGGATACTTGAATCGAAA 300
DB 6049 GTTCTCTCCGCTTATCTATATGCGAAGAAATTCCTGGGGCGGATACTTGAATCGAAA 6108
QY 301 CTTTGGGCGGGGCAAGTAGTGTCTCGTTGCCATTCGACTTGGACACCATTTGGATG 360
DB 6109 CTTTGGGCGGGGCAAGTAGTGTCTCGTTGCCATTCGACTTGGACACCATTTGGATG 6168
QY 361 GATTTCTCCGAAGCAACTATATCAAGCCCTGTTTTTGTGAGCGGTAAGATGTGCA 420
DB 6169 GATTTCTCCGAAGCAACTATATCAAGCCCTGTTTTTGTGAGCGGTAAGATGTGCA 6228
QY 421 CCGATGATCTTAGTCAATTCGTTGCGGCAATTCCTAAGACTGCGGGCTTTTCCGAAC 480
DB 6229 CCGATGATCTTAGTCAATTCGTTGCGGCAATTCCTAAGACTGCGGGCTTTTCCGAAC 6288
QY 481 CTGCGCAATGCGCTGTACGAGAAATGCGATGAATGCGTTACCGGGTTTACCATACC 540
DB 6289 CTGCGCAATGCGCTGTACGAGAAATGCGATGAATGCGTTACCGGGTTTACCATACC 6348
QY 541 CTTGAAAGGGGCGGTGCCATTTGACATGTAAGCTTATGCTGAAAACCTGATGCTGAAGGGT 600
DB 6349 CTTGAAAGGGGCGGTGCCATTTGACATGTAAGCTTATGCTGAAAACCTGATGCTGAAGGGT 6408
QY 601 TCGGCAAGTTCTTTTCCAACTGCACTGCTACGACTGCGAAGCCGTTTTTTGACCAA 660
DB 6409 TCGGCAAGTTCTTTTCCAACTGCACTGCTACGACTGCGAAGCCGTTTTTTGACCAA 6468
QY 661 TGTTCGATAGTAGAGCGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 720
DB 6469 TGTTCGATAGTAGAGCGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 6528
QY 721 GTCAATTGGCGTGCATTTTCGGACTGCTGCTGCGAAGCAAGCTGTTCAATGCTGGGTA 780
DB 6529 GTCAATTGGCGTGCATTTTCGGACTGCTGCTGCGAAGCAAGCTGTTCAATGCTGGGTA 6588
QY 781 GACGATGTTACATATATGAAAGCAAGTAGTCTGTGAGGCAAGCTTTGGTCAATGCT 840
DB 6589 GACGATGTTACATATATGAAAGCAAGTAGTCTGTGAGGCAAGCTTTGGTCAATGCT 6648
QY 841 TTCAGGAGACCTCTAGTGTCTGTGCGAATGGGAGCGATGCGATTTCTTCTGCTGCA 900
DB 6649 TTCAGGAGACCTCTAGTGTCTGTGCGAATGGGAGCGATGCGATTTCTTCTGCTGCA 6708
QY 901 TTTCTGCTGTTTTTCTTCTGAGCGTTACGGCTGTCTTCAATGAGGCGGTTCCCAAT 960
DB 6709 TTTCTGCTGTTTTTCTTCTGAGCGTTACGGCTGTCTTCAATGAGGCGGTTCCCAAT 6768
QY 961 CCGGCGCACTTCACACTTACTGTTACTACCAAGGCGTCCAAATACATGTGAAAAGCCGGG 1020
DB 6769 CCGGCGCACTTCACACTTACTGTTACTACCAAGGCGTCCAAATACATGTGAAAAGCCGGG 6828
QY 1021 CAGCTGCCACCGAAGCTGTTCATCGCGTTTACACAGGTTGGCGGTTCTTGAAGGAC 1080
DB 6829 CAGCTGCCACCGAAGCTGTTCATCGCGTTTACACAGGTTGGCGGTTCTTGAAGGAC 6888
QY 1081 GGTTCATGAGCGAGATATGTTGGCTTCCGCTGTGCTATATCAAGGCGCTTGAAA 1140
DB 6889 GGTTCATGAGCGAGATATGTTGGCTTCCGCTGTGCTATATCAAGGCGCTTGAAA 6948

QY 1141 TCAGGACACTAGTAGTGCGCTCATGACTCTCGGCAAAATTTGGCTGAACCGTTTCGGGAG 1200
DB 6949 TCAGGACACTAGTAGTGCGCTCATGACTCTCGGCAAAATTTGGCTGAACCGTTTCGGGAG 7008
QY 1201 GAGTCTCTCTCTTCAAGGATAGAGAGATCTTTCTGGGCAACATCTCTCTGGTGT-AA 1259
DB 7009 GAGTCTCTCTCTTCAAGGATAGAGAGATCTTTCTGGGCAACATCTCTCTGGTGTGAA 7068
QY 1260 ACATGAGTTTTTCTCATGATTTGGGACCTTATCAAGCTAAATGGGAATGGAATCTGGCGG 1319
DB 7069 ACATGAGTTTTTCTCATGATTTGGGACCTTATCAAGCTAAATGGGAATGGAATCTGGC-GG 7127
QY 1320 GTTTGGTCCAGTTTTTGAAGCGGGTTTATGAGATCTCTCGCTGGTCAATCAAGGATA 1379
DB 7128 GTTTGGTCCAGTTTTTGAAGCGGGTTTATGAGATCTCTCGCTGGTCAATCAAGGATA 7187
QY 1380 TGAAGAAATCAGCGGATGTCCCTGAAGGAATCTGAGAACTTTCACAGTGGATTCGATC 1439
DB 7188 TGAAGAAATCAGCGGATGTCCCTGAAGGAATCTGAGAACTTTCACAGTGGATTCGATC 7247
QY 1440 TGAAGTGTTAACGGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAAGCGGAT 1499
DB 7248 TGAAGTGTTAACGGTGTGTGTGTGAGCCAGCGCATATGCCATGTTCAAGTCAAGCGGAT 7307
QY 1500 TCAGAGGAAAAAGACAAAAATAAAGATTAAGGCTTAAGAGCGGGATATCGAATTTATGA 1559
DB 7308 TCAGAGGAAAAAGACAAAAATAAAGATTAAGGCTTAAGAGCGGGATATCGAATTTATGA 7367
QY 1560 TAAAGTGTGTGTCACTGTGACTGCGAAATATCCAACTCAGGACTTGCCTGACATGCGA 1619
DB 7368 TAAAGTGTGTGTCACTGTGACTGCGAAATATCCAACTCAGGACTTGCCTGACATGCGA 7427
QY 1620 TACCAATTTTTTCAAGCACCACTGAAACCAAGCGTTGATTAACGCCATATGACAGATC 1679
DB 7428 TACCAATTTTTTCAAGCACCACTGAAACCAAGCGTTGATTAACGCCATATGACAGATC 7487
QY 1680 GTCAAAACTCTTCTCATGACTGGAAGGAAATTCGCTTGAACCATATCTCCGCTTG 1739
DB 7488 GTCAAAACTCTTCTCATGACTGGAAGGAAATTCGCTTGAACCATATCTCCGCTTG 7547
QY 1740 TGTCTCTATGAGCGGATCGCAAAAGCAGTGTATTCCTGAGCTATAGATGCGAGATCC 1799
DB 7548 TGTCTCTATGAGCGGATCGCAAAAGCAGTGTATTCCTGAGCTATAGATGCGAGATCC 7607
QY 1800 GAAATGTTAAAGTCTATGCTCATCACTTATACATGGAAGAGGACTCCCAAGCTGTT 1859
DB 7608 GAAATGTTAAAGTCTATGCTCATCACTTATACATGGAAGAGGACTCCCAAGCTGTT 7667
QY 1860 GCGGCTCCCGCAAAAAGAGCGATTATGCTGCTGCGGAGCGCAATTCGAGATCTTT 1919
DB 7668 GCGGCTCCCGCAAAAAGAGCGATTATGCTGCTGCGGAGCGCAATTCGAGATCTTT 7727
QY 1920 CCGGCGGTTTGCACACACTATTTCTGCTGCTGCTGATTAAGCAAAATGTTATTTCA 1979
DB 7728 CCGGCGGTTTGCACACACTATTTCTGCTGCTGCTGATTAAGCAAAATGTTATTTCA 7787
QY 1980 ACATGATGCTTACAGAGAGATGCGGGGAGGCTTTCAAACTCAACCGGCTGTGA 2039
DB 7788 ACATGATGCTTACAGAGAGATGCGGGGAGGCTTTCAAACTCAACCGGCTGTGA 7847
QY 2040 GGAATTTTATTCGAGAACTTTTCTTCAAGCACTGAGACAGGCTTAATGATACCGGAGT 2099
DB 7848 GGAATTTTATTCGAGAACTTTTCTTCAAGCACTGAGACAGGCTTAATGATACCGGAGT 7907
QY 2100 TTAATTTGGGCGGTTGAGTGTCTTCTTCAAGGATGAGTGGAGGCTGTTATGAGAC 2159
DB 7908 TTAATTTGGGCGGTTGAGTGTCTTCTTCAAGGATGAGTGGAGGCTGTTATGAGAC 7967
QY 2160 CCGGTGTAACCGCTGTGTGCAATTAATCAATATGGAAGGCAATTTGGCAAAAGGCA 2219
DB 7968 CCGGTGTAACCGCTGTGTGCAATTAATCAATATGGAAGGCAATTTGGCAAAAGGCA 8027
QY 2220 TCTCTCGAACTCTTGGAGAGATATACTAACCGCAATTAAG 2268

DB 8028 TCCTCTGGAACACTCTTGGAAGATATTACTACCGACGACTAGAAATTAG 8076

RESULT 4
AAV69756
ID AAV69756 standard; DNA; 2211 BP.
XX AAV69756;
AC
DT 01-MAR-1999 (first entry)
XX
DE Trypophan oxygenase laam gene.
XX
KM Trypophan oxygenase; laam gene; auxin; indoleacetic acid;
XX transgenic plant; seedless fruit; tomato; watermelon; cucumber; ds.
OS Agrobacterium tumefaciens.
XX
FH Key Location/Qualifiers
FT CDS 4..2205
/*tag= a
XX
PD 12-NOV-1998.
XX
XX MO9849888-A1.
XX
XX 06-MAY-1998; 98MO-US09013.
XX
XX 06-MAY-1997; 97US-0045725.
XX
XX (UNIV) UNIV KANSAS STATE RES FOUND.
XX
XX L1 Y;
XX
XX MPI; 1999-034673/03.
XX
XX P-PSDB; AAV61574.
XX
XX A new construct to express phytohormones in developing fruit -
XX useful for, e.g. producing substantially seedless fruit from
XX transgenic plants
XX
XX Example 1; Page 21-26; 49pp; English.
XX
XX This is the nucleotide sequence of the Agrobacterium tumefaciens
XX laam gene that codes for tryptophan oxygenase (see AAV61574). This
XX enzyme catalyzes the conversion of tryptophan to indoleacetamide
XX in the biosynthetic pathway of the auxin indoleacetic acid. A
XX claimed DNA construct comprises either an isopentenyl transferase
XX (see AAV61575) or a tryptophan oxygenase encoding sequence, operably
XX linked to an ovary or developing fruit-specific plant-expressible
XX promoter (see AAV69755 and AAV69759). The construct is used to stably
XX integrate enzymes involved in cytokinin or auxin biosynthesis into
XX the plant genome to achieve a transgenic plant (preferably tomato,
XX cucumber or watermelon) producing seedless fruit in the absence of
XX pollination.
XX
XX Sequence 2211 BP; 528 A; 504 C; 595 G; 584 T; 0 other;
XX
XX Query Match 95.4%; Score 2162.8; DB 20; Length 2211;
XX Best Local Similarity 99.4%; Pred. No. 0;
XX Matches 2192; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

QY 64 ACAATGTCGATAGGCGGATGAATTGACCGCAGGGTTCGATGCTCTTGAACGA 123
DB 1 ACAATGTCGATAGGCGGATGAATTGACCGCAGGGTTCGATGCTCTTGAACGA 60
QY 124 GAAGCTTCTAGGGAGAGAGATTACTCAAGCTCCACCGAGTCCAGCGCTGGTAACT 183
DB 61 GAAGCTTCTAGGGAGAGAGATTACTCAAGCTCCACCGAGTCCAGCGCTGGTAACT 120
QY 184 TGCAAAAGGCTGGCGATGCTGCTCCCGAGATCTCAGCTGATGAAAGGTAGCAGTT 243

DB 121 TGCAAAAGGCTGGCGATGCTGCTCCCGAGATCTCAGCTGATGAAAGGTAGCAGTT 180
QY 244 CTCTCCGCTTATATCTATATTTGGCAAGAAATTTCTGGGCGGATTAATTGAATGAAACCT 303
DB 181 CTCTCCGCTTATATCTATATTTGGCAAGAAATTTCTGGGCGGATTAATTGAATGAAACCT 240
QY 304 TGGGCGGCGCAACAGTGAAGTGTCTGTTGCGCATGAGCTTGGAACAACATTCATGAT 363
DB 241 TGGGCGGCGCAACAGTGAAGTGTCTGTTGCGCATGAGCTTGGAACAACATTCATGAT 300
QY 364 TTCTCCGAAGCAACAACTAATCCAAAGCCCTGTTTCTGAGCGGTAAAGATGCAACCG 423
DB 301 TTCTCCGAAGCAACAACTAATCCAAAGCCCTGTTTCTGAGCGGTAAAGATGCAACCG 360
QY 424 ATTGATCTTATGATATTTGCTGGCATTTTCAATCTCTAAGCTGCGGCTTTGCAACCTG 483
DB 361 ATTGATCTTATGATATTTGCTGGCATTTTCAATCTCTAAGCTGCGGCTTTGCAACCTG 420
QY 484 CCAATGCGGCTGACAGAAATGGCAGATGAATAATGCGTTACCGGTTTACATAACCTT 543
DB 421 CCAATGCGGCTGACAGAAATGGCAGATGAATAATGCGTTACCGGTTTACATAACCTT 480
QY 544 GAAAGGCGCGTGCATTTGACATGATGCTTATGTCGAAACCTGATGCTGAAGGCTTCG 603
DB 481 GAAAGGCGCGTGCATTTGACATGATGCTTATGTCGAAACCTGATGCTGAAGGCTTCG 540
QY 604 GCAGGTTCTTTTCCAACTAATCGATGCTCTCAAGCTGACAGACCGTTTATGCAATGT 663
DB 541 GCAGGTTCTTTTCCAACTAATCGATGCTCTCAAGCTGACAGACCGTTTATGCAATGT 600
QY 664 TCCGATAGTGAAGAGATGCGCTCTTCCGGAAGATGTTCTTAAGCGGAAGTGGCGGTC 723
DB 601 TCCGATAGTGAAGAGATGCGCTCTTCCGGAAGATGTTCTTAAGCGGAAGTGGCGGTC 660
QY 724 ATTGGCGCTGCAATTTCCGGAAGTGTGTCGCAACGAACCTGCTTACGTCGGGTAGAC 783
DB 661 ATTGGCGCTGCAATTTCCGGAAGTGTGTCGCAACGAACCTGCTTACGTCGGGTAGAC 720
QY 784 GATGTTCAATATATGAGCAATGATGCTGTTGAGGCAAGCTTTGTCACATGCTTTC 843
DB 721 GATGTTCAATATATGAGCAATGATGCTGTTGAGGCAAGCTTTGTCACATGCTTTC 780
QY 844 AGGAGCGCTCTTATGTCGTCGTCGGAATGGGGGCGATGCGATTTCTCTGTCGATTC 903
DB 781 AGGAGCGCTCTTATGTCGTCGTCGGAATGGGGGCGATGCGATTTCTCTGTCGATTC 840
QY 904 TGGTGTGTTTCTTCTCTGAGCGTTACGCGCTGTCTTCGATGAGCGCTTCCAAATCCC 963
DB 841 TGGTGTGTTTCTTCTCTGAGCGTTACGCGCTGTCTTCGATGAGCGCTTCCAAATCCC 900
QY 964 GGCACAGTGCACCTTACTTGTGTCACCAAGGCTCCAAATCAATGTGGAAGCGCGGACG 1023
DB 901 GGCACAGTGCACCTTACTTGTGTCACCAAGGCTCCAAATCAATGTGGAAGCGCGGACG 960
QY 1024 CTGCGACCGAAGCTGTTCCATGCGCTTACCAAGCTTGGCGCTGTTCTTGAAGACGCT 1083
DB 961 CTGCGACCGAAGCTGTTCCATGCGCTTACCAAGCTTGGCGCTGTTCTTGAAGACGCT 1020
QY 1084 TTTCATGAGCGAGATATGTTGTTGCTGCTGCTGCTGCTATTAATCAAGGCTTGAATCA 1143
DB 1021 TTTCATGAGCGAGATATGTTGTTGCTGCTGCTGCTGCTATTAATCAAGGCTTGAATCA 1080
QY 1144 GGCACATTTAGTGGGCTCATGCTCTGCGCAAAATTTGGCTGAACCGTTTCGGGAGGAG 1203
DB 1081 GGCACATTTAGTGGGCTCATGCTCTGCGCAAAATTTGGCTGAACCGTTTCGGGAGGAG 1140
QY 1204 TCCTTCTCTCAGAGATAGAGAGATCTTCTGGGCAACATCTCTGCTGGTGT- AAACA 1262
DB 1141 TCCTTCTCTCAGAGATAGAGAGATCTTCTGGGCAACATCTCTGCTGGTGTGTAACA 1200
QY 1263 TGAGATTTTCTCATGATATGGAACCTATTCAGCTAATGGAATAGAGATGCGGGGTT 1322
DB 1201 TGAGATTTTCTCATGATATGGAACCTATTCAGCTAATGGAATAGAGATGCGG- GGGTT 1259

```

QY 1323 TGGTCCAGTTTGTGAAGCGGGTTATTGAGATCTCCGGTGGTCAACAAGATATGA 1382
Db 1260 TGGTCCAGTTTGTGAAGCGGGTTATTGAGATCTCCGGTGGTCAACAAGATATGA 1319
QY 1383 AGAAATACAGCGGATGTCCTTGAAGAAATCTCAAACTTCCAGTCCGATCTGA 1442
Db 1320 AGAAATACAGCGGATGTCCTTGAAGAAATCTCAAACTTCCAGTCCGATCTGA 1379
QY 1443 AGTGGTTAAGGTTGTCTGTGAGCCGCGCATATGCGCATTTCAATCAGGCGATTC 1502
Db 1380 AGTGGTTAAGGTTGTCTGTGAGCCGCGCATATGCGCATTTCAATCAGGCGATTC 1439
QY 1503 GAAGAAAAACAAAAATTAAGATTAAGCTTAAGAGCGGGATATCTGAATTTATGATA 1562
Db 1440 GAAGAAAAACAAAAATTAAGATTAAGCTTAAGAGCGGGATATCTGAATTTATGATA 1499
QY 1563 GGTGGTGTCACTTGTGACTCGCAATATCCAACTGAGCATTTGCTGACATGCGATAC 1622
Db 1500 GGTGGTGTCACTTGTGACTCGCAATATCCAACTGAGCATTTGCTGACATGCGATAC 1559
QY 1623 CAATATTTTTCAGGCAACCAATGAAACGAGGTTGATTAACGCGATATGACAGATGTC 1682
Db 1560 CAATATTTTTCAGGCAACCAATGAAACGAGGTTGATTAACGCGATATGACAGATGTC 1619
QY 1683 AAAACTCTTCTGTATGATGAAAGAAATCTGGTTAGACATATCTCCGCTTGTGT 1742
Db 1620 AAAACTCTTCTGTATGATGAAAGAAATCTGGTTAGACATATCTCCGCTTGTGT 1679
QY 1743 CCTCATGAGAGGGAATGCAAAAGCAGTATGCTGCTGATGAGTGCAGAGATCCGAA 1802
Db 1680 CCTCATGAGAGGGAATGCAAAAGCAGTATGCTGCTGATGAGTGCAGAGATCCGAA 1739
QY 1803 TGGTAAAGGTTCTAGTGTCTCATCACTTATCATGAGAGAGATCTCCCAAGCTTTGGC 1862
Db 1740 TGGTAAAGGTTCTAGTGTCTCATCACTTATCATGAGAGAGATCTCCCAAGCTTTGGC 1799
QY 1863 GGTCCCGACAAAAGAGGATTAATCTGCTGCGGAGCGCAATTCAGATCTTTCCC 1922
Db 1800 GGTCCCGACAAAAGAGGATTAATCTGCTGCGGAGCGCAATTCAGATCTTTCCC 1859
QY 1923 GGGGTTTCCGACAGCACTATTTCTGCTGCTGCTGATTAAGCAAAATGTTATTACA 1982
Db 1860 GGGGTTTCCGACAGCACTATTTCTGCTGCTGCTGATTAAGCAAAATGTTATTACA 1919
QY 1983 TGAATGGCTTACAGACGAGATGCGGGGAGCTTTCAAACTCAACCGGCGTGTAGGA 2042
Db 1920 TGAATGGCTTACAGACGAGATGCGGGGAGCTTTCAAACTCAACCGGCGTGTAGGA 1979
QY 2043 TTTTAAATTCGAAGAACTTTCTTTCAAGCACTGAGCAAGGCTAATGATCCGGAATTA 2102
Db 1980 TTTTAAATTCGAAGAACTTTCTTTCAAGCACTGAGCAAGGCTAATGATCCGGAATTA 2039
QY 2103 CTGGGCGGGTGGAGTGTCTTCAACAGTGGATGGGTGAAGGCTATTGACACCG 2162
Db 2040 CTGGGCGGGTGGAGTGTCTTCAACAGTGGATGGGTGAAGGCTATTGACACCG 2099
QY 2163 GTGTAAAGCGGCTGTGCAATTAATCCAAATGTGAGGCAATTTGGCAAAAGGCAATCC 2222
Db 2100 GTGTAAAGCGGCTGTGCAATTAATCCAAATGTGAGGCAATTTGGCAAAAGGCAATCC 2159
QY 2223 TCTCGAAGCACTTGGAGAGATTAATCAACCGCAGTAAAGAAATTAAG 2288
Db 2160 TCTCGAAGCACTTGGAGAGATTAATCAACCGCAGTAAAGAAATTAAG 2205

```

RESULT 5
 AAT48868/c
 ID AAT48868 standard; DNA; 6112 BP.
 AC AAT48868;
 XX
 XX 07-APR-1997 (first entry)
 DT

```

XX Iams expression vector pHP5838.
DE Male sterile plant; hybrid seed; pollen; microspore; oilseed;
XX canola; rape; antisense RNA; stress resistance;
KW herbicide resistance; transgenic plant; crop protection;
KW Iams; indole acetamide synthase; Bp10; promoter; Brassica;
KM pHP5838; de; cyclic.
XX
OS Chimeric Brassica sp.;
OS Chimeric Agrobacterium tumefaciens.
XX
FH Key Location/Qualifiers
FT misc_difference 3794
FT FT /tag= a
FT FT /note= "base 3794 is given as v in the
FT specification"
XX
PD MO9640950-AI.
XX
PN 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US08692.
XX
PR 07-JUN-1995; 95US-0476864.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albani DJ, Arnison PG, Fabijanski SF, Laurian RS;
XX WPI; 1997-052340/05.
XX
PT Recombinant molecule for preparation of male sterile plants - used
PT to produce stress resistant plants and for increase in seed yield
XX
PS Example 39a; Fig 27a-e; 224pp; English.
XX
CC Plasmid pHP5838 (AAT48868) contains the Brassica microspore-specific
CC Bp10 promoter linked to the indole acetamide synthase (Iams) gene
CC from the Agrobacterium tumefaciens Ti plasmid derivative pPCV311,
CC and to the nos 3' non-translated region. The Iams gene is involved
CC in tumour formation in Agrobacterium-infected plants. Plasmids
CC pHP5838-41 (see also AAT48869-71), contg. Iams or Iamh genes under
CC control of anther-specific Bp10 or TA39 promoters, were introduced
CC into canola lines by transformation. Male sterile plants were
CC obt'd. These can be used in the prodn. of hybrid seed using pollen
CC from a male fertile line.
XX
SQ Sequence 6112 BP; 1515 A; 1466 C; 1451 G; 1679 T; 1 other;
Query Match 90.0%; Score 2040.8; DB 18; Length 6112;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2139; Conservative 0; Mismatches 127; Indels 2; Gaps 2;
QY 1 ATGTGAGCTTCACTCTCTGATTAACAGTGGATCATCTCCCAACAAATGGTGGAT 60
Db 2870 ATGTGAGCTTCACTCTCTGATTAACAGTGGATCATCTCTCAACAAATGGTGGAT 2811
QY 61 CTGACATAGTGGATTAAGCGGATGAATTGACCGCAGGTTTCCGATCCTTCTAGAA 120
Db 2810 CTGATATAGTGGATTAAGCGGATGAATTGACCGCAGGTTTCCGATCCTTCTAGAA 2751
QY 121 CGAAGAGCTTCTAGGGGAAGAGATTAATCAAGCTCCACCGAGTGCAGCGCTGGTTA 180
Db 2750 CGTGAAGCTTCTAGGGGAAGAGATTAATCAAGCTCCCGGAGTGCAGCGCTGGTTA 2691
QY 181 GCTTGCAAAAGGCTGGCGATGGTGGCTTCCCGAGATCTCAGCTGTGAAAGATAGCA 240
Db 2690 GCTTGCAAAAGGCTGGCGAGTGGCTTCCCGAGATCTCAGCTGTGAAAGATAGCA 2631
QY 241 GTTCTCTCCGCTTAATATCTAATTTGGCAAGAAATTTCTGGGCGGATATCTTGATGAA 300
Db 2630 GCGCTCTCCGCTTAATATCTAATTTGGCAAGAAATTTCTGGGCGGATATCTTGATGAA 2571

```


QY 301 CCTTGGGCGGGGCAACAGTGAAGTGTCTGTTGCCATCGACTTGAACACCATTTGATG 360
 Db 2570 CTTTGGGCGGGAGAGAGAGTGAAGTGTCTGTTGCCATCGACTTGAACACCATTTGATG 2511
 QY 361 GATTTCCTCGAAGACACACTTAATCCAAAGCCCTGTTTGTCTGAGCGGTAAGATGTGA 420
 Db 2510 GATTTCCTCGAAGACACACTTAATCCAAAGCCCTGTTTGTCTGAGCGGTAAGATGTGA 2451
 QY 421 CCGATTGATCTTGTGATCTTGTGAGCATTTCAATCTTAAGATGCGCGCTTTGCAAC 480
 Db 2450 TCAAGCATTTGTGATCTTGTGAGCATTTCAATCTTAAGATGCGCGCTTTGCAAC 2391
 QY 481 CTGCAATGCGCGCTTGAAGAAATGCGACATGAATGCGTTACCGGGTTTACCAATACC 540
 Db 2390 CTGCAATGCGCGCTTGAAGAAATGCGACATGAATGCGTTACCGGGTTTACCAATACC 2331
 QY 541 CTGGAAGGGGCGGTCCTGATTTGACATGTAGCTTATGTGCGAACTGATGCTGAAGGT 600
 Db 2330 CTGGAAGGGGCGGTCCTGATTTGACATGTAGCTTATGTGCGAACTGATGCTGAAGGT 2271
 QY 601 TCGGACAGTTCCTTTCCAAACAATGCAATGCTCTACAGCTGCAAGCGTTTGTGACAA 660
 Db 2270 TCGGACAGTTCCTTTCCAAACAATGCAATGCTCTACAGCTGCAAGCGTTTGTGACAA 2211
 QY 661 TGTTCGATAGTGAAGGATCGGCTCTTTCCGAGAGATGTTCTTAAGCCGAAAGTGCG 720
 Db 2210 TGTTCGATAGTGAAGGATCGGCTCTTTCCGAGAGATGTTCTTAAGCCGAAAGTGCG 2151
 QY 721 GTCATTTGGGCTGGCATTTCCGACCTGCTGAGCGAAACGAACTGCTTCAATGCTGGGTA 780
 Db 2150 ATCATTTGGGCTGGCATTTCCGACCTGCTGAGCGAAACGAACTGCTTCAATGCTGGGTA 2091
 QY 781 GAGAGTTTCAATATATGAAGCAAGATGCTGTTGAGAGCAAGCTTGTGACATGCT 840
 Db 2090 GAGAGTTTCAATATATGAAGCAAGATGCTGTTGAGAGCAAGCTTGTGACATGCT 2031
 QY 841 TTGAGGAGCGCTCTAGTGTCTGAGCGGAAATGAGGCGGATGCGATTTCTCTGCTGCA 900
 Db 2030 TTGAGGAGCGCTCTAGTGTCTGAGCGGAAATGAGGCGGATGCGATTTCTCTGCTGCA 1971
 QY 901 TTTGCTTGTGTTTCTTCTGAGCGGCTGTTCTTGTGATGAGGCGGTTCCAAAT 960
 Db 1970 TTTGCTTGTGTTTCTTCTGAGCGGCTGTTCTTGTGATGAGGCGGTTCCAAAT 1911
 QY 961 CCGGCGCACTGCACTTAATGCTGTTACCAAGCGCTCCAAATGATGAGGCGG 1020
 Db 1910 CCGGCGCACTGCACTTAATGCTGTTACCAAGCGCTCCAAATGATGAGGCGG 1851
 QY 1021 CAGGTGCGACCGAAGCTGTTCCATCGGTTTCAACAGGTTGCGGCTTCTGAAGGAC 1080
 Db 1850 CAGGTGCGACCGAAGCTGTTCCATCGGTTTCAACAGGTTGCGGCTTCTGAAGGAC 1791
 QY 1081 GATTTCATGAGCAGATATTTGTTGGCTTCCGCTGCTGATTAATCAAGCCTTGA 1140
 Db 1790 GATTTCATGAGCAGATATTTGTTGGCTTCCGCTGCTGATTAATCAAGCCTTGA 1731
 QY 1141 TCAAGACACATTAAGTGGGCTGATGATCTCTGCGTAAATTTGGTGAACCGTTTGGGAGG 1200
 Db 1730 TCAAGACACATTAAGTGGGCTGATGATCTCTGCGTAAATTTGGTGAACCGTTTGGGAGG 1671
 QY 1201 GAGTCTTCTCTTCAAGGATAGAGAGATCTTCTGAGGACACATCTCTCGTGGTGT -AA 1259
 Db 1670 GAGTCTTCTCTTCAAGGATAGAGAGATCTTCTGAGGACACATCTCTCGTGGTGTAA 1611
 QY 1260 ACATGAGTTTCTCTCATGATTTGGACCTATTCAAGCTAAATGGGAATAGATCTGGCGG 1319
 Db 1610 ACATGAGTTTCTCTCATGATTTGGACCTATTCAAGCTAAATGGGAATAGATCTGGCGG -GG 1552
 QY 1320 GTTTGCTCAGTTTGTGAAGCGGTTTATTTGAATCTCTCGCTTGTGATCAACGAGATA 1379
 Db 1551 GTTTGCTCAGTTTGTGAAGCGGTTTATTTGAATCTCTCGCTTGTGATCAACGAGATA 1492

QY 1380 TGAAGAAAAACAGCGGATGTGCCCTGAAGAAATCTGAGAACTTCCAGCTGGATCGCATC 1439
 Db 1491 TGAAGAAAAACAGCGGATGTGCCCTGAAGAAATCTGAGAACTTCCAGCTGGATCGCATC 1432
 QY 1440 TGAAGTGTGAACGCTGTGTCTGTGAGCCAGCCATATGCTCATGTTCAAGTCAAGGCGAT 1499
 Db 1431 TGAAGTGTGAACGCTGTGTCTGTGAGCCAGCCATATGCTCATGTTCAAGTCAAGGCGAT 1372
 QY 1500 TCAAGAAAAAGACAAAAATTAAGATTAAGGCTTAAGAGCGGGATATCTGAATTTATGA 1559
 Db 1371 TGAAGAAAAAGACAAAAATTAAGATTAAGGCTTAAGAGCGGGATATCTGAATTTATGA 1312
 QY 1560 TAAAGTGTGTGATCATCTGCACTGCAATATCAACTCAAGCATTTGCTGATGCA 1619
 Db 1311 TAAAGTGTGTGATCATCTGCACTGCAATATCAACTCAAGCATTTGCTGATGCA 1252
 QY 1620 TACCAATTTTTTCAAGCACCACTGAACCAAGCGGTTGATTAACAGCCATATGACAGATC 1679
 Db 1251 TACCAATTTTTTCAAGCACCACTGAACCAAGCGGTTGATTAACAGCCATATGACAGATC 1192
 QY 1680 GTCAAAAGCTTCTGATGATGATGCAAGAAATCTGTTAGACATATCTCCGCTTTG 1739
 Db 1191 GTCAAAAGCTTCTGATGATGATGCAAGAAATCTGTTAGACATATCTCCGCTTTG 1132
 QY 1740 TGTCTCATGAGCGGATGCAAAAGCAGTATGCTGAGCTATGATGAGGAGATCC 1799
 Db 1131 TGTCTCATGAGCGGATGCAAAAGCAGTATGCTGAGCTATGATGAGGAGATCC 1072
 QY 1800 GAATGTTAAAGTCTAGTGTCTCATGATTTATCATGAGAGAGCATCTCCACAAAGCTGT 1859
 Db 1071 GAATGTTAAAGTCTAGTGTCTCATGATTTATCATGAGAGAGCATCTCCACAAAGCTGT 1012
 QY 1860 GGGGCTCCCGACAAAAGAGGCAATATGCTGCTGCGGAGCGCAATTTGAGATCTTT 1919
 Db 1011 GGGGCTCCCGACAAAAGAGGCAATATGCTGCTGCGGAGCGCAATTTGAGATCTTT 952
 QY 1920 CCGGCGGTTTCCAGCACCTATTTCTGCTGCGGCTGATTTACGACCAAAATGTTATCA 1979
 Db 951 CCGGCGGTTTCCAGCACCTATTTCTGCTGCGGCTGATTTACGACCAAAATGTTATCA 892
 QY 1980 ACATGATTTGCTTACAGAGAAATGCGGGGAGCTTTCAACTCAACCGGCTGTGA 2039
 Db 891 ACATGATTTGCTTACAGAGAAATGCGGGGAGCTTTCAACTCAACCGGCTGTGA 832
 QY 2040 GGAATTTTATTTGAAGAACTTTCTTTCAAGCACTGACACCGGCTAATGATACCGGAGT 2099
 Db 831 GGAATTTTATTTGAAGAACTTTCTTTCAAGCGCTGAGACATGATATACCGGAGT 772
 QY 2100 TTAATTGGCGGATGCAATTTCTTCAACAGGTGATGGGTGAGGAGGCTATTCAGAC 2159
 Db 771 TTAATTGGCGGATGCAATTTCTTCAACAGGTGATGGGTGAGGAGGCTATTCAGAC 712
 QY 2160 CCGGTGTAACGCGCTGTGTCATTTATCCCAATTTGTGAGGAGATTTTGGCAAGGCGAA 2219
 Db 711 CCGGTGTAACGCGCTGTGTCATTTATCCCAATTTGTGAGGAGATTTTGGCAAGGCGAA 652
 QY 2220 TCCCTCGAACACTCTTGAAGAGATATACTACCGCAGTAAATTA 2267
 Db 651 TCCCTCGAACACTCTTGAAGAGATATACTACCGCAGTAAATTA 604

RESULT 6
 AAT4870/c
 ID AAT4870 standard; DNA; 7208 BP.
 XX AAT4870;
 AC
 XX
 DT 07-APR-1997 (first entry)
 XX
 DE 1amh expression vector pHP5840.
 XX
 KM Male sterile plant; hybrid seed; pollen; microspore; oilseed;
 KM canola; rape; antisense RNA; stress resistance;

KM herbicide resistance; transgenic plant; crop protection;
 KM iamh; indole acetamide hydrolase; T339; promoter; tobacco;
 KM pHP5840; ds; cyclic.
 XX
 OS Chimeric Nicotiana tabacum;
 OS Chimeric Agrobacterium tumefaciens.
 XX
 PH Key Location/Qualifiers
 FT misc_difference 4890
 FT /tag= a
 FT /note= "base 4890 is given as v in the
 specification"
 FT
 PN W09640950-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 07-JUN-1996; 96WO-US08692.
 XX
 PR 07-JUN-1995; 95US-0476864.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 PI Albani DJ, Arnison PG, Fabijanski SF, Laurian RS;
 XX
 DR WPI; 1997-052340/05.
 XX
 PT Recombinant molecule for preparation of male sterile plants - used
 XX to produce stress resistant plants and for increase in seed yield
 XX
 PS Example 39a; Fig 29a-f; 224pp; English.
 XX
 CC Plasmid pHP5840 (AAT48870) contains the tobacco anther-specific
 CC T339(BB3) promoter and the o' nontranslated leader operably linked
 CC to the indole acetamide hydrolase (iamh) gene from the Agrobacterium
 CC tumefaciens Ti plasmid derivative pPCV311, and to the P-II 3'
 CC nontranslated region. The iamh gene is involved in tumor
 CC formation in Agrobacterium-infected plants. Plasmid pHP5838-41
 CC (see also AAT48868-71), contg. iamS or iamH genes under control of
 CC anther-specific Bp10 or T339 promoters, were introduced into canola
 CC lines by transformation. Male sterile plants were obtd. These can
 CC be used in the prodn. of hybrid seed using pollen from a male
 CC fertile line.
 CC
 XX
 SQ Sequence 7208 BP; 1938 A; 1619 C; 1640 G; 2010 T; 1 other;
 Query Match 90.0%; Score 2040.8; DB 18; Length 7208;
 Best Local Similarity 94.3%; Pred. No. 0;
 Matches 2139; Conservative 0; Mismatches 127; Indels 2; Gaps 2;
 QY 1 ATGTCAGCTTCACCTCTCTTGATACCAAGCGATCATCTCCCAACCAAAATGATGAT 60
 DB 2924 ATGTCAGCTTCACCTCTCTTGATACCAAGCGATCATCTCTACCAAAATGATGAT 2865
 QY 61 CTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 DB 2864 CTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2805
 QY 121 CGAAGACCTTCTAGGAGAGAGATTAACAAGCTCAACGAGTCAGCGCTGAGTTA 180
 DB 2804 CGTGAACCTTCTAGGAGAGAGATTAACAAGCTCAACGAGTCAGCGCTGAGTTA 2745
 QY 181 GCTTGAAGAGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 DB 2744 GCTTGAAGAGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2685
 QY 241 GTTCTCTCCCTTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 300
 DB 2684 GCCTCTCCCTTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 2625
 QY 301 CTTTGGGCGCGGCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 DB 2624 CTTTGGGCGCGGCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2565

QY 361 GATTTCCGAGACACAACTATCCAAGCCGTTTTGCTGAGCGGTAAAGATGTGCA 420
 DB 2564 GATTTCCGAGACACAACTATCCAAGCCGTTTTGCTGAGCGGTAAAGATGTGCA 2505
 QY 421 CCGATGATCTTATGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 DB 2504 TCAGAGATCTTATGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2445
 QY 481 CTGCAATGCGCTGTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 DB 2444 CTGCAATGCGCTGTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2385
 QY 541 CTTGAAGGCGCGTGCATTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 2384 CTTGAAGGCGCGTGCATTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2325
 QY 601 TCGGAGGTTCTTTCCAACAATCGACTTCTACGACTGACAGACCGTTTTTGAACCA 660
 DB 2324 TCGGAGGTTCTTTCCAACAATCGACTTCTACGACTGACAGATCGTTTTTGAACCA 2265
 QY 661 TGTTCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 DB 2264 TGTTCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2205
 QY 721 GTGATTTGGGCTGGGCAATTTCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 DB 2204 ATATTGGGCTGGGCAATTTCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2145
 QY 781 GACGATTTACATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 DB 2144 GACGATTTACATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2085
 QY 841 TTGAGGAGGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 DB 2084 TTGAGGAGGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2025
 QY 901 TTTGCTGTTTTTTCTTCTCCAGCGCTTACGCGCTTCTTCAATGAGCGCTTCCCAAT 960
 DB 2024 TTTGCTGTTTTTTCTTCTCCAGCGCTTACGCGCTTCTTCAATGAGCGCTTCCCAAT 1965
 QY 961 CCGGCAAGTGCACACTTACTTGTCTTACCAAGCGCTTCAATGATGATGATGATGATGATGATGAT 1020
 DB 1964 CCGGCAAGTGCACACTTACTTGTCTTACCAAGCGCTTCAATGATGATGATGATGATGATGATGAT 1905
 QY 1021 CAGCTGCCACCGAAGCTTTCATGCTGCTTCAACGCTTGGCGTCTTTGAAGAC 1080
 DB 1904 CAGCTGCCACCGAAGCTTTCATGCTGCTTCAACGCTTGGCGTCTTTGAAGAC 1845
 QY 1081 GCTTTTCATGAGGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 DB 1844 GCTTTTCATGAGGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1785
 QY 1141 TCAGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 1784 TCAGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1725
 QY 1201 GAGTCTTCTTTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
 DB 1724 GAGTCTTCTTTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1665
 QY 1260 ACATGAGATTTTCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319
 DB 1664 ACATGAGATTTTCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1606
 QY 1320 GTTGTGTCAGTTTGAAGGCGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1379
 DB 1605 GTTGTGTCAGTTTGAAGGCGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1546
 QY 1380 TGAAGAAATACAGCGGATGCTTGAAGAAATTCAGAACTTTCAGCTGAGATGCAATC 1439
 DB 1545 TGAAGAAATACAGCGGATGCTTGAAGAAATTCAGAACTTTCAGCTGAGATGCAATC 1486

QY 1440 TGAAGTGTAAACGCTGTCTGTGAGCCAGCGCATATGCGCATGTTCAGTCAAGGCGCAT 1499
DB 1485 TCAAGTGTAAACGCTGTCTGTGAGCCAGCGCATATGCGCATGTTCAGTCAAGGCGCAT 1426
QY 1500 TCGAAGGAAAAGCAAAAATTAAGATTAAGCTTAAGAGCGGATATCTGAATCTTATGA 1559
DB 1425 TGAAGAAAGAAAACAAAAATTAAGATTAAGAGCTTAAGAGCGGATATCTGAATCTTATGA 1366
QY 1560 TAAAGTGTGTGATCAATCTGAGTCCGCAATATCAATCAATGAGGATGCGTCAATGCA 1519
DB 1365 TAAAGTGTGTGATCAATCTGAGTCCGCAATATCAATCAATGAGGATGCGTCAATGCA 1306
QY 1620 TACCAATATTTTTCAGGACACAGTGAACCAAGCGGTTGATTAACGCCATATGAAGGATC 1679
DB 1305 TACCAATATTTTTCAGGACACAGTGAACCAAGCGGTTGATTAACGCCATATGAAGGATC 1246
QY 1680 GTCAAAACTCTTCTGTATGATGATGACCAAAATTTCTGTTAGCATATCTCCGCTTTG 1739
DB 1245 GTCAAAACTCTTCTGTATGATGATGACCAAAATTTCTGTTAGCATATCTCCGCTTTG 1186
QY 1740 TGTCTCATGAGAGGATGCGCAAAAGAGTGTATTTGCTGAGTATGAGTCCGAGATCC 1799
DB 1185 TGTCTCATGAGAGGATGCGCAAAAGAGTGTATTTGCTGAGTATGAGTCCGAGATCC 1126
QY 1800 GAATGTAAAGTCTAGTGTCTCATAGTATTAACAAGGAGAGGATCTCCGCAAGCTGT 1859
DB 1125 GAATGTAAAGTCTAGTGTCTCATAGTATTAACAAGGAGAGGATCTCCGCAAGCTGT 1066
QY 1860 GGGGGTCCCCGACAAAAGAGGATATGCTGCTGCGGAGCGCAATTTTCAGATCTTT 1919
DB 1065 GGGGGTCCCCGACAAAAGAGGATATGCTGCTGCGGAGCGCAATTTTCAGATCTTT 1006
QY 1920 CCGGCGCTTGGCCAGGACCATTTCTGCTGCGGCTGATTAAGACCAAAATGTTATTC 1979
DB 1005 CCGGCGCTTGGCCAGGACCATTTCTGCTGCGGCTGATTAAGACCAAAATGTTATTC 946
QY 1980 ACATGATTTGCTTAACAGACGAAATGCGGGGAGCTTTCAAACTCAACCGCGTGTGA 2039
DB 945 ACATGATTTGCTTAACAGACGAAATGCGGGGAGCTTTCAAACTCAACCGCGTGTGA 886
QY 2040 GGAATTTTATTTCTGAAGAACTTTTCTTCAAGCACTGAGACGCGTAAATGATCCGAGT 2099
DB 885 GGAATTTTATTTCTGAAGAACTTTTCTTCAAGCGCTGAGATGATGATATATCCGAGT 826
QY 2100 TTAAGTGGGAGTGTGAGTGTCTTCAAGAGTGAATGAGGAGGCTGATTTAGAC 2159
DB 825 TTAAGTGGGAGTGTGAGTGTCTTCAAGAGTGAATGAGGAGGCTGATTTAGAC 766
QY 2160 CCGGCTGTAAAGCGCTGTGTCAATTTATCACAATTTGTGAAGCAATTTTGGCAAGGCGCA 2219
DB 765 CCGGCTGTAAAGCGCTGTGTCAATTTATCACAATTTGTGAAGCAATTTTGGCAAGGCGCA 706
QY 2220 TCTCTCTGAACACTCTTGAAGAGATTAATCAATCCGAGTGAATTA 2267
DB 705 TCTCTCTGAACACTCTTGAAGAGATTAATCAATCCGAGTGAATTA 658

RESULT 7
AADD0626
ID AADD0626 standard; DNA; 2544 BP.
XX
AC AADD0626;
XX
DT 29-AUG-2000 (first entry)
XX
DE Agrobacterium tumefaciens ipt/iaam untranslatable BR construct.
XX
KM iaam; iaah; ipt; T-DNA; transferred DNA; gall-disease resistance; BR;
KW bacterial resistance construct; untranslatable RNA; tumour; oncogene;
XX plant growth hormone; de.
XX
OS Agrobacterium tumefaciens.
XX

FH Key Location/Qualifiers
FT misc_feature 1..309
FT /tag= a
FT /note= "Corresponds to Agrobacterium mutant ipt gene"
FT misc_feature 303..2105
FT /tag= b
FT /note= "Corresponds to Agrobacterium mutant iaam gene"
XX
XX MO200026346-A1.
XX
PD 11-MAY-2000.
XX
XX 04-NOV-1999; 99MO-US26100.
XX
XX 05-NOV-1998; 98US-0107185.
XX
XX (UYOR-) UNIV OREGON STATE.
XX
XX Ream W, Mok MC, Lee H;
XX
XX WPI; 2000-36599/31.
XX
PT Gall resistant plants generated by transforming cells with an
PT untranslatable nucleic acid homologous to a gall disease-causing gene
XX
XX Example c; Page 44-45; 49pp; English.
XX
CC The patent discloses a method for producing plants resistant to gall
CC disease, induced by Agrobacterium infection, by transformation with a
CC bacterial resistance (BR) construct capable of eliciting co-suppression.
CC The construct encodes an untranslatable RNA molecule, that is highly
CC homologous to the tumour or gall disease-causing genes iaam, iaah and
CC ipt of Agrobacterium tumefaciens. These oncogenes are modified by
CC introduction of premature termination codons or frameshift mutations, to
CC inhibit their expression. This method is used to reduce susceptibility
CC of plants to gall diseases. The present DNA sequence is the Agrobacterium
CC tumefaciens bacterial resistance construct, that encodes untranslatable
CC RNA forms of both the mutant ipt and iaam genes. This construct can
CC suppress both shooty and necrotic gall formation in transgenic plants.
XX
SQ Sequence 2544 BP; 625 A; 592 C; 674 G; 653 T; 0 other;
XX
Query Match 78.9%; Score 1788.6; DB 21; Length 2544;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1791; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 12 ACCTCTCTTGAATTAACGATGCGATCATCTCCCAACCAAAATGTGATTCGATGCT 71
DB 318 ACCTCTCTTGAATTAACGATGCGATCATCTCCCAACCAAAATGTGATTCGATGCT 377
QY 72 CGATTAAGGCGGATGAATTTGGAACCGAGGGTTCCGATGCTTTTGAACGAGAGCTTC 131
DB 378 CGATTAAGGCGGATGAATTTGGAACCGAGGGTTCCGATGCTTTTGAACGAGAGCTTC 437
QY 132 TAGGGAAGGAGGATTAATCAAGCTCAACGAGTCAAGGCTGGTGAAGCTTGAAG 191
DB 438 TAGGGAAGGAGGATTAATCAAGCTCAACGAGTCAAGGCTGGTGAAGCTTGAAG 497
QY 192 GGTGGCCGATGTGCTTCCCGAGATCTGAGTGTGGAAGGTAGCACTTCTCCGC 251
DB 498 GGTGGCCGATGTGCTTCCCGAGATCTGAGTGTGGAAGGTAGCACTTCTCCGC 557
QY 252 TTAATCTATATTTGGAAGAAATTTCTGGGGGCGATTAATGCAAACTTGGGGGCG 311
DB 558 TTAATCTATATTTGGAAGAAATTTCTGGGGGCGATTAATGCAAACTTGGGGGCG 617
QY 312 GGCACAGTGAAGTGTCTGCTGCTCAATGCACTTGAACACCATTTGATGATTTCCGA 371
DB 618 GGCACAGTGAAGTGTCTGCTGCTCAATGCACTTGAACACCATTTGATGATTTCCGA 677
QY 372 AGCAACAATTAATCAAGCCCTGTTTGTGCTGACGCGTAAAGATGACCGATGATCT 431

Db 678 AGCAACAATAATCCAGCCCTGTTTGTGAGGCGTAAAGATGTGACCCGATTGATCT 737
Qy 432 TAGTCATTTGCTGGCCATTTCAATCTCTAAGACTGCCGGCTTTGGAAACCCCTGCAATGCC 491
Db 738 TAGTCATTTGCTGGCCATTTCAATCTCTAAGACTGCCGGCTTTGGAAACCCCTGCAATGCC 797
Qy 492 GCTGTACGAGATGAGCAAGATGAATGCGTTACCGGGTTTACATTAACCCCTTGAAGGGGC 551
Db 798 GCTGTACGAGATGAGCAAGATGAATGCGTTACCGGGTTTACATTAACCCCTTGAAGGGGC 857
Qy 552 CGTGTACATTTGACATGTGATCTTATGCTCGAAACCTGATGCTGAAGGGTTGCGCAGTTTC 611
Db 858 CGTGTACATTTGACATGTGATCTTATGCTCGAAACCTGATGCTGAAGGGTTGCGCAGTTTC 917
Qy 612 CTTTCCAAACATGCACTTGTCTACGACTGCGAGCCGTTTTTTTGAACCAATGTTCCGATAG 671
Db 918 CTTTCCAAACATGCACTTGTCTACGACTGCGAGCCGTTTTTTTGAACCAATGTTCCGATAG 977
Qy 672 TGAACGATCGGCTCTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCGGTCAATTGGCGC 731
Db 978 TGAACGATCGGCTCTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCGGTCAATTGGCGC 1037
Qy 732 TGGCATTTCCGAGACTCGTGTGCGAAACGAACTGCTTATGCTGGGTGACGATGTTAC 791
Db 1038 TGGCATTTCCGAGACTCGTGTGCGAAACGAACTGCTTATGCTGGGTGACGATGTTAC 1097
Qy 792 AATATGAAGCAAGTATGCTGTGAGGCAACCTTGGCAATGCTTCCGAGGACGC 851
Db 1098 AATATGAAGCAAGTATGCTGTGAGGCAACCTTGGCAATGCTTCCGAGGACGC 1157
Qy 852 TCCATGATGCTGTGCGCCGAAATGGGGGCGATGCGATTTCTCTGCTGCAATTCCTGCTGT 911
Db 1158 TCCATGATGCTGTGCGCCGAAATGGGGGCGATGCGATTTCTCTGCTGCAATTCCTGCTGT 1217
Qy 912 TTTTCTTCTGAGCGCTTACCGGCTGTCTTGTGATGAGCGCTTCCAAATCCCGCACAGT 971
Db 1218 TTTTCTTCTGAGCGCTTACCGGCTGTCTTGTGATGAGCGCTTCCAAATCCCGCACAGT 1277
Qy 972 CGACACTTACTGTGCTACCAAGGCGTCCAAATATATGGAAGAACCGGCGACGCTGCCACC 1031
Db 1278 CGACACTTACTGTGCTACCAAGGCGTCCAAATATATGGAAGAACCGGCGACGCTGCCACC 1337
Qy 1032 GAAGCTTTCATCGCGCTTACCAAGGCTGTGCGCTGCTTGTGAAGGACGGTTTTCATGA 1091
Db 1338 GAAGCTTTCATCGCGCTTACCAAGGCTGTGCGCTGCTTGTGAAGGACGGTTTTCATGA 1397
Qy 1092 GCGAGATATTGTGTGCTTGCCTGTGCTATTAATCAAGCCCTTGAATCAGGACACAT 1151
Db 1398 GCGAGATATTGTGTGCTTGCCTGTGCTATTAATCAAGCCCTTGAATCAGGACACAT 1457
Qy 1152 TAGGTGGCTTCATGATCTCTGCGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1211
Db 1458 TAGGTGGCTTCATGATCTCTGCGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1517
Qy 1212 TTCAGGATAGAGAGATCTTCTTGGGACACATCCTCGTGTGATGAACATGAGATTTT 1271
Db 1518 TTCAGGATAGAGAGATCTTCTTGGGACACATCCTCGTGTGATGAACATGAGATTTT 1577
Qy 1272 CTTATGATTGGGACTTATTAAGCTATGGAATAGGATCTGGCGGGTTTGGTCCAGT 1331
Db 1578 CTTATGATTGGGACTTATTAAGCTATGGAATAGGATCTGGCGGGTTTGGTCCAGT 1637
Qy 1332 TTTTGAAGGGGGTTTATGAGATCTCTCGCTTGTGTATCAAGGATATGAAGAAATCA 1391
Db 1638 TTTTGAAGGGGGTTTATGAGATCTCTCGCTTGTGTATCAAGGATATGAAGAAATCA 1697
Qy 1392 GCGGATGCTCCCTGAAGGAATCTCAGAACTTCCAGTCCGATCGCATCTGAAGGTTAA 1451
Db 1698 GCGGATGCTCCCTGAAGGAATCTCAGAACTTCCAGTCCGATCGCATCTGAAGGTTAA 1757
Qy 1452 CCGTGTGTCTGTGAGCCAGCGCATATGCAATGTTCAAGTCAGGGCGAATTCAGAGGAAA 1511
Db 1758 CCGTGTGTCTGTGAGCCAGCGCATATGCAATGTTCAAGTCAGGGCGAATTCAGAGGAAA 1817

Qy 1512 GACAAAATTAAGATTAAGCTTAAGACCGGATATCTGAACCTTTATGATTAAGGTGGT 1571
Db 1818 GACAAAATTAAGATTAAGCTTAAGACCGGATATCTGAACCTTTATGATTAAGGTGGT 1877
Qy 1572 CACATCTGACTGCAAAATATCCAACTCAAGCATTCCTGACATGCGATACCAATATTTT 1631
Db 1878 CACATCTGACTGCAAAATATCCAACTCAAGCATTCCTGACATGCGATACCAATATTTT 1937
Qy 1632 TCAGGACACAGTAAACCAAGCGTTGATTAACAGCCATATGACAGATTCCTCAAACTCTT 1691
Db 1938 TCAGGACACAGTAAACCAAGCGTTGATTAACAGCCATATGACAGATTCCTCAAACTCTT 1997
Qy 1692 CTTGATGACTGAACGAAAATTTCTGTTAAGCATATCTCCGCTTGTGTCTCATAGA 1751
Db 1998 CTTGATGACTGAACGAAAATTTCTGTTAAGCATATCTCCGCTTGTGTCTCATAGA 2057
Qy 1752 CCGGATCGCAAAAGCAGTGTATTCCTGACATATGAGTCGACAGATCCGAATGCT 1806
Db 2058 CCGGATCGCAAAAGCAGTGTATTCCTGACATATGAGTCGACAGATCCCACTGCT 2112

RESULT 8
AAD00625
ID AAD00625 standard; DNA; 1807 BP.
XX
AC AAD00625;
XX
DT 29-AUG-2000 (first entry)
XX
DE Agrobacterium tumefaciens T-DNA mutant iaam gene.
XX
KW iaam; iaah; ipc; T-DNA; transferred DNA; gall-disease resistance; BR;
KW bacterial resistance construct; untranslatable RNA; tumour; oncogene;
KW auxin; plant growth hormone; mutant; ds.
XX
OS Agrobacterium tumefaciens.
XX
PN W0200026346-A1.
XX
PD 11-MAY-2000.
XX
PF 04-NOV-1999; 99MO-US26100.
XX
PR 05-NOV-1998; 98US-0107185.
XX
PA (UYOR-) UNIV OREGON STATE.
XX
PI Ream W, Mok MC, Lee H;
XX
DR WPI; 2000-36559/31.
XX
PT Gall resistant plants generated by transforming cells with an
PT untranslatable nucleic acid homologous to a gall disease-causing gene
PT -
PS Example c; Page 43-44; 49pp; English.
XX
CC The patent discloses a method for producing plants resistant to gall
CC disease, induced by Agrobacterium infection, by transformation with a
CC bacterial resistance (BR) construct capable of eliciting co-suppression.
CC The construct encodes an untranslatable RNA molecule, that is highly
CC homologous to the tumour or gall disease-causing genes iaam, iaah and
CC ipt of Agrobacterium tumefaciens. These oncogenes are modified by
CC introduction of premature termination codons or frameshift mutations, to
CC inhibit their expression. This method is used to reduce susceptibility
CC of plants to gall diseases. The present DNA sequence is the Agrobacterium
CC tumefaciens T-DNA (transferred DNA) mutant iaam gene. This oncogene was
CC modified by introduction of a stop sequence (TGA) at the third codon and
CC deletion of two bases following the third codon. This sequence was used
CC for the construction of double stranded RNA and untranslatable dsRNA
CC vectors, that reduce gall formation in transformed plant cells.
XX

Sequence 1807 BP; 428 A; 418 C; 486 G; 475 T; 0 other;

Query Match 78.8%; Score 1788; DB 21; Length 1807;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 12 ACCTCTCTGATTAACCGATGCGATCATCTCCCAACCAAAATGATGATCGAATGCT 71
DB 18 ACCTCTCTGATTAACCGATGCGATCATCTCCCAACCAAAATGATGATCGAATGCT 77
QY 72 CGATTAAGCGCGATGAATTTGGACCGCGAGGTTTCCGATGCTTCTTAAGAGAAAGTTT 131
DB 78 CGATTAAGCGCGATGAATTTGGACCGCGAGGTTTCCGATGCTTCTTAAGAGAAAGTTT 137
QY 132 TAGGGAGAGAGGATTAATCAAAAGCTCACCGAGTGCAGGCTGGGTAGCTTGCAGAAAG 191
DB 138 TAGGGAGAGAGGATTAATCAAAAGCTCACCGAGTGCAGGCTGGGTAGCTTGCAGAAAG 197
QY 192 GCTGGCCGATGTCGCTTCCCGAGATCTGAGCTGTGAAAGAGTAGCACTTCTCCGC 251
DB 198 GCTGGCCGATGTCGCTTCCCGAGATCTGAGCTGTGAAAGAGTAGCACTTCTCCGC 257
QY 252 TTATATCTATTTGGCAAAGAAATTTGGGGCGGATCTTGAATCGAACTTGGGGCG 311
DB 258 TTATATCTATTTGGCAAAGAAATTTGGGGCGGATCTTGAATCGAACTTGGGGCG 317
QY 312 GGCAGAGTAGAGTGTCTGTTGCGATCGAATTGGAACCACTTGTGATTTCTCGA 371
DB 318 GGCAGAGTAGAGTGTCTGTTGCGATCGAATTGGAACCACTTGTGATTTCTCGA 377
QY 372 AGCACAATAATCAAGCCCTGTTTTTGTGAGCGGTAAGAAAGATGTCACCGATGATCT 431
DB 378 AGCACAATAATCAAGCCCTGTTTTTGTGAGCGGTAAGAAAGATGTCACCGATGATCT 437
QY 432 TAGTCATTTGTCGCCATTTCAATCTCTAAGACTGCCGCTTTGAAACCTCGCAATGCC 491
DB 438 TAGTCATTTGTCGCCATTTCAATCTCTAAGACTGCCGCTTTGAAACCTCGCAATGCC 497
QY 492 GCTGTAGAGAAATGGCAGATGAAATGCGTTACCGGGTTTACCAATAACCTTGAAGGGG 551
DB 498 GCTGTAGAGAAATGGCAGATGAAATGCGTTACCGGGTTTACCAATAACCTTGAAGGGG 557
QY 552 CGTGCCATTTGACATGCTAGCTTATGCTGAAACCTGATGCTGAAGGGTTCCGCAAGTTT 611
DB 558 CGTGCCATTTGACATGCTAGCTTATGCTGAAACCTGATGCTGAAGGGTTCCGCAAGTTT 617
QY 612 CTTTCCAAATCGACTTGTCTTAAGACTGCAACCGTTTTTTTGAACCAATGTTCCGATG 671
DB 618 CTTTCCAAATCGACTTGTCTTAAGACTGCAACCGTTTTTTTGAACCAATGTTCCGATG 677
QY 672 TGGACGAGATGGGCTTCTTCCGAGGATGTTCCCTAAGCCGAAAGTGCGGGCATTTGGCG 731
DB 678 TGGACGAGATGGGCTTCTTCCGAGGATGTTCCCTAAGCCGAAAGTGCGGGCATTTGGCG 737
QY 732 TGGCATTTCCGAGCTCGTGTGCGAAACGAATGCTTCAATGCTGGGGTAGACGATGTTAC 791
DB 738 TGGCATTTCCGAGCTCGTGTGCGAAACGAATGCTTCAATGCTGGGGTAGACGATGTTAC 797
QY 792 AATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTTGGTCAATGCTTTCAAGGACGC 851
DB 798 AATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTTGGTCAATGCTTTCAAGGACGC 857
QY 852 TCCAGTGTGCTGGCGGAAATGGGGGGGATGCGATTTCTCTGCTGCGATTTGCTGCTGTT 911
DB 858 TCCAGTGTGCTGGCGGAAATGGGGGGGATGCGATTTCTCTGCTGCGATTTGCTGCTGTT 917
QY 912 TTTCTTCTCGAGCGTTACGCGCTGTCTTGCATGAGGCGGTTCCCAATCCGCGACAGT 971
DB 918 TTTCTTCTCGAGCGTTACGCGCGTGTCTTGCATGAGGCGGTTCCCAATCCGCGACAGT 977
QY 972 CGACACTTAATCTGTTACCAAGGCTCAATATCATGTGAAAGCGGGGAGCTGCCACC 1031
DB 978 CGACACTTAATCTGTTACCAAGGCTCAATATCATGTGAAAGCGGGGAGCTGCCACC 1037

```

```

QY 1032 GAAGCTGTTCCATCGCGTTTACCAACCGTTGGCGTGGCTTCTTGAAGACGGTTTTCATGA 1091
DB 1038 GAAGCTGTTCCATCGCGTTTACCAACCGTTGGCGTGGCTTCTTGAAGACGGTTTTCATGA 1097
QY 1092 GCGAGATTAATTTGTTGGCTTCCGCTTCCGCTATTAATCAAGGCTTTGAATCAGACACAT 1151
DB 1098 GCGAGATTAATTTGTTGGCTTCCGCTTCCGCTATTAATCAAGGCTTTGAATCAGACACAT 1157
QY 1152 TAGTGGGCTCATGATCCCGGCAATTTGGGCAACCGTTTGGGAGGGAGTCTTCTC 1211
DB 1158 TAGTGGGCTCATGATCCCGGCAATTTGGGCAACCGTTTGGGAGGGAGTCTTCTC 1217
QY 1212 TTCAGGATTAAGAGGATCTTCTTGGGCAACACTCTCTGTTGATTAACATGAGTCTT 1271
DB 1218 TTCAGGATTAAGAGGATCTTCTTGGGCAACACTCTCTGTTGATTAACATGAGTCTT 1277
QY 1272 CCTCATGATTTGGGACCTATTAAGCTTAATGGAATAGATCTTGGCGGGTTTGGTCCAGT 1331
DB 1278 CCTCATGATTTGGGACCTATTAAGCTTAATGGAATAGATCTTGGCGGGTTTGGTCCAGT 1337
QY 1332 TTTTGAAGGGGTTTATGAGATCTCTCGGCTTGGTCAATCAACGATATGAAGAAATCA 1391
DB 1338 TTTTGAAGGGGTTTATGAGATCTCTCGGCTTGGTCAATCAACGATATGAAGAAATCA 1397
QY 1392 GCGGATGTCCCTGAAGAAATCTCAGAACTTCCACGTCGAGATGCAATCTGAACTGTTAA 1451
DB 1398 GCGGATGTCCCTGAAGAAATCTCAGAACTTCCACGTCGAGATGCAATCTGAACTGTTAA 1457
QY 1452 CGGTGTGCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAAGGCGCATTCAGAAAGAAA 1511
DB 1458 CGGTGTGCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAAGGCGCATTCAGAAAGAAA 1517
QY 1512 GACAAAATTAAGGTTAAGGCTTAAGAGCGGAGATATCGAACTTATGATTAAGGTGGTGT 1571
DB 1518 GACAAAATTAAGGTTAAGGCTTAAGAGCGGAGATATCGAACTTATGATTAAGGTGGTGT 1577
QY 1572 CACATCTGACCTCGCAATATCCAACTCAGGCACTTCCGTCAGATGCGATATCAATATTTT 1631
DB 1578 CACATCTGACCTCGCAATATCCAACTCAGGCACTTCCGTCAGATGCGATATCAATATTTT 1637
QY 1632 TCAGGACCAAGTGAACCAAGCGGTTATTAACAGCCATATGACAGATTCCTCAAACTCTT 1691
DB 1638 TCAGGACCAAGTGAACCAAGCGGTTATTAACAGCCATATGACAGATTCCTCAAACTCTT 1697
QY 1692 CCGATGACCTGAACGAAATTCGTTAGACATATCCTCCGCTTGTGTCTCATGGA 1751
DB 1698 CCGATGACCTGAACGAAATTCGTTAGACATATCCTCCGCTTGTGTCTCATGGA 1757
QY 1752 CGGATTCGAAAAGCAGTATTCCTGACATATGAGTGCAGAGATCC 1799
DB 1758 CGGATTCGAAAAGCAGTATTCCTGACATATGAGTGCAGAGATCC 1805

```

RESULT 9
AAN60970
ID AAN60970 standard; DNA; 2126 BP.

AC AAN60970;
XX 25-MAR-2003 (updated)
DT 28-OCT-1991 (first entry)
XX
DE Tl-DNA region of A. rhizogene agropine-type plasmid pRiHRI.
XX
KW Ri plasmid; plant plasmid.
XX
OS Agrobacterium rhizogenes HRI and A4.
PN EP204590-A.
XX
PD 10-DEC-1986.
XX

XX (PEKE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX P-PSDB; ABB69563.
DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 35480; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4734 BP; 1394 A; 978 C; 1048 G; 1314 T; 0 other;
Query Match 1.8%; Score 39.8; DB 23; Length 4734;
Best Local Similarity 52.0%; Pred. No. 0.28;
Matches 89; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 690 TCCGAGAGATGTTCTTAAGCCGAAAGTGGCGGTATGGCGCTGCTTCCGACTCT 749
DB 1219 TTTCAGGAATATATACCAACGTTAAGATCGATGATGAGCTGGATGGCTGCTATC 1278
QY 750 GGTGGCAACGAACTGCTTCAATGCTGGGATAGACATGTTACATATATGAAAGCAATGA 809
DB 1279 GGCTGCAATACCTTTTGCAAAAGCGCTGTGACGATTTCTTATTTCTGAGGACGGGG 1338
QY 810 TCGTGTGAGGCAAGCTTTGTCACATGCTTTTCAGGAGCGCTCTAGTGT 860
DB 1339 CCGGGTGGGGGAGCGCATTTGTGTCATACCGCTAGCAATATCAGAAAGT 1389
RESULT 11
AB15662
ID ABL15662 standard; cDNA; 4800 BP.
XX
AC ABL15662;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 41468.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX

DR WPI; 2001-656860/75.
DR P-PSDB; ABB71559.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 41468; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4800 BP; 1265 A; 1116 C; 1141 G; 1278 T; 0 other;
Query Match 1.8%; Score 39.8; DB 23; Length 4800;
Best Local Similarity 52.0%; Pred. No. 0.28;
Matches 89; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 690 TCCGAGAGATGTTCTTAAGCCGAAAGTGGCGGTATGGCGCTGCTTCCGACTCT 749
DB 4496 TTTCAGGAATATATACCAACGTTAAGATCGATGATGAGCTGGATGGCTGCTATC 4555
QY 750 GGTGGCAACGAACTGCTTCAATGCTGGGATAGACATGTTACATATATGAAAGCAATGA 809
DB 4556 GGCTGCAATACCTTTTGCAAAAGCGCTGTGACGATTTCTTATTTCTGAGGACGGGG 4615
QY 810 TCGTGTGAGGCAAGCTTTGTCACATGCTTTTCAGGAGCGCTCTAGTGT 860
DB 4616 CCGGGTGGGGGAGCGCATTTGTGTCATACCGCTAGCAATATCAGAAAGT 4666
RESULT 12
AB13667
ID ABL13667 standard; cDNA; 1872 BP.
XX
AC ABL13667;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 35483.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR P-PSDB; ABB69564.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -
XX
PS Claim 1; SEQ ID NO 35483; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB16511), expressed DNA
CC sequences (AB57737-AB57739) and the encoded proteins
CC (AB57737-AB57739).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1872 BP; 499 A; 477 C; 524 G; 372 T; 0 other;
XX
Query Match 1.7%; Score 39.2; DB 23; Length 1872;
Best Local Similarity 52.4%; Pred. No. 0.25; Mismatches 0; Gaps 0;
Matches 86; Conservative 0; Indels 78; Indels 0; Gaps 0;
XX
QY 690 TCCGAGAGATGTTCTTAAGCGGAGGCGGCTGATGCGCTGATTCGGACTCGT 749
DB 219 TTCAGGGAATTAATCAACGTTAAGATCGATCATTTGAGCTGAGATGCGCTGATC 278
XX
QY 750 GGTGGCAACGAACTGCTTCACTGCTGGGCTAGAGATGTTACATATATGAAGCAAGTGA 809
DB 279 GGCTGCAAAATCACCTTTTGCAAAACGCTGTGAGATTTCTTATTCGTGAGGACGCGG 338
XX
QY 810 TCGGTGTGAGAGCAAGCTTTGTCACATGCTTTCAGGAGCGTC 853
DB 339 CCGGCTGGCGGAGCGCATTTGTCATACCGCTCAGCAATATATC 382
XX
RESULT 13
AAK15122
ID AAK15122 standard; cDNA to mRNA; 1575 BP.
XX
AC AAK15122;
XX
DT 23-APR-1999 (first entry)
XX
DE Nucleic acid encoding an apoptosis inducing protein.
XX
KW Cell death; apoptosis; inhibition; proliferation; cancer cell;
KW apoptosis inducing protein; AIP; chub mackerel; anticancer; ss.
XX
OS Scomber japonicus.
XX
FH Key
FT CDS Location/Qualifiers
FT 1..1575
FT /*tag= a
XX
PN MO9852972-A1.
XX
PD 26-NOV-1998.
XX
PF 22-MAY-1998; 98MO-JF02261.
XX
PR 23-MAY-1997; 97JP-0133549.
XX
PA (TENS-) TENSBI SUIBAN CO LTD.
XX
PI Iwamoto M, Jung S;
XX
DR WPI; 1999-070139/06.
XX
DR P-PDB; AAM96805.
XX
PT Apoptosis inducing protein from mackerel and gene encoding it - for
XX use as anticancer agents
XX
PS Claim 5; Page 32-34; 47pp; Japanese.

XX
CC The present sequence encodes a protein which induces cell death
CC (apoptosis) and inhibits the proliferation of cancer cells. The
CC protein (apoptosis inducing protein, AIP) is isolated from chub
CC mackerel. The protein can be used as an anticancer agent and as a
CC reagent for study of the mechanisms of apoptosis in vitro.
XX
SQ Sequence 1575 BP; 436 A; 358 C; 406 G; 375 T; 0 other;
XX
Query Match 1.7%; Score 38.6; DB 20; Length 1575;
Best Local Similarity 66.1%; Pred. No. 0.35;
Matches 72; Conservative 0; Mismatches 34; Indels 3; Gaps 1;
XX
QY 713 AAGTGGCGGATGCGCTGCGATTTCCGAGTCTGTGCGAAGCAACTGCTCATG 772
DB 182 ATGTGGTTATAGTCGAGCTGGCATGCCGAGTACGCGCGCAATGTCGCAAGACG 241
XX
QY 773 CTGGGCTAGACGATGTTACATATATGAAGCAAGTATGTTGAGG 821
DB 242 CAGG---ACACACGATACCATTTGGAGGCTATATGCTGTTGAGG 287
XX
RESULT 14
ABZ56743
ID ABZ56743 standard; cDNA; 468 BP.
XX
AC ABZ56743;
XX
DT 28-MAR-2003 (first entry)
XX
DE *Aspergillus oryzae* polynucleotide SEQ ID NO 5856.
XX
KW *Aspergillus oryzae*; fermentation; fungus; industrial; EST;
KW expressed sequence tag; gene; ss.
XX
OS *Aspergillus oryzae*.
XX
PN WO200279476-A1.
XX
PD 10-OCT-2002.
XX
PF 22-MAR-2002; 2002MO-IB00890.
XX
PR 30-MAR-2001; 2001JP-0098371.
XX
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (NARE-) NAT RES INST BREWING.
PA (NORQ) NAT FOOD RES INST MIN AGRIC.
XX
PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;
PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
XX
DR WPI; 2003-046817/04.
XX
PT Detection of expression of specific *Aspergillus* genes for monitoring
PT the fermentation and growth conditions of the fungus, using DNA probes
XX
PS Claim 1; SEQ ID NO 5856; 48pp + Sequence Listing; Japanese.
XX
CC The invention relates to a polynucleotide having any of 6006 specific
CC sequences (ABZ50888-ABZ58893), which are expressed by a fungus under
CC specific culture conditions including one or more of eutrophic,
CC oligotrophic, solid, early germination, alkaline, high temperature, low
CC temperature or maltose culture or polynucleotides stringently hybridizing
CC to these sequences. The polynucleotides are useful for monitoring the
CC progress of fermentation and the growth conditions of a fungus,
CC especially of *Aspergillus oryzae* which is widely used in industrial
CC fermentation. Also monitoring for fungal contamination.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 20:47:41 ; Search time 8367.46 Seconds
(without alignments)
11088.547 Million cell updates/sec

Title: US-09-434-837-10

Perfect score: 2268

Sequence: 1 atgcacagcttcacactctctcct.....actacgcagtagaattag 2268

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: GenBank: *
2: gb_ba: *
3: gb_hlg: *
4: gb_in: *
5: gb_om: *
6: gb_ov: *
7: gb_pat: *
8: gb_ph: *
9: gb_pl: *
10: gb_pr: *
11: gb_ro: *
12: gb_sts: *
13: gb_sy: *
14: gb_un: *
15: gb_vi: *
16: em_ba: *
17: em_fun: *
18: em_hum: *
19: em_in: *
20: em_mu: *
21: em_om: *
22: em_or: *
23: em_ov: *
24: em_pat: *
25: em_ph: *
26: em_pl: *
27: em_ro: *
28: em_sts: *
29: em_un: *
30: em_vi: *
31: em_hlg_hum: *
32: em_hlg_inv: *
33: em_hlg_mus: *
34: em_hlg_pia: *
35: em_hlg_rtd: *
36: em_hlg_mam: *
37: em_hlg_vtc: *
38: em_sy: *
39: em_higo_hum: *
40: em_higo_mus: *
41: em_higo_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2230.6	98.4	24595	1 ATACH5	X00493 Agrobacteri
2	2230.6	98.4	24595	6 BD016312	BD016312 Method of
3	2230.6	98.4	24595	6 E00404	E00404 Ti plasmid
4	2230.6	98.4	24595	6 E00546	E00546 DNA fragment
5	2230.6	98.4	194140	1 AF242881	AF242881 Agrobacte
6	2225.8	98.1	2591	1 TIMPS2	K02554 Ti plasmid
7	2044	90.1	9931	1 AE007927	AE007927 Agrobacte
8	2044	90.1	10377	1 AE009419	AE009419 Agrobacte
9	2044	90.1	206479	1 AB016260	AB016260 Agrobacte
10	2042.4	89.6	4660	1 AF126446	AF126446 Agrobacte
11	2032.8	89.6	4660	1 AB025110	AB025110 Agrobacte
12	2016.2	88.9	29524	1 ATTU37588	ATTU37588 Agrobacte
13	1880.8	82.9	10200	1 ATTH1A8H	X56185 A. tumefacie
14	1880.8	82.9	15463	1 AV083987	U83987 Agrobacteri
15	1880.8	82.9	15463	1 AV083986	U83986 Agrobacteri
16	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
17	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
18	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
19	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
20	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
21	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
22	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
23	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
24	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
25	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
26	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
27	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
28	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
29	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
30	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
31	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
32	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
33	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
34	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
35	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
36	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
37	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
38	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
39	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
40	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
41	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
42	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
43	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
44	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte
45	1880.8	82.9	14960	1 AF126447	AF126447 Agrobacte

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
1	ATACH5	Agrobacterium tumefaciens Ti plasmid pTi15955-T-DNA region.	X00493	J05108	X00282	Agrobacterium tumefaciens	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.	1 (bases 1 to 24595)	Barker, R.F., Idler, K.B., Thompson, D.V., and Kemp, J.D.	Nucleotide sequence of the T-DNA region from the Agrobacterium

```
REFERENCE JOURNAL tumefaciens octopine Ti plasmid pTI15955
AUTHORS Plant Mol. Biol. 2, 335-350 (1983)
TITLE 2 (bases 602 to 14237)
JOURNAL Gleen,U., De Beuckeleer,M., Seurinck,J., Deboeck,F., De Greve,H.,
MEDLINE Lemmerys,M., Van Montagu,M., and Schell,J.
PUBMED The complete nucleotide sequence of the Ti-DNA of the Agrobacterium
REFERENCE tumefaciens plasmid pTiAch5
AUTHORS EMBO J. 3 (4), 835-846 (1984)
TITLE 6327292
JOURNAL 3 (bases 1 to 24595)
MEDLINE Turk,S.C., Nester,E.W. and Hooykaas,P.J.
PUBMED The virA promoter is a host-range determinant in Agrobacterium
REFERENCE tumefaciens
AUTHORS Mol. Microbiol. 7 (5), 719-724 (1993)
TITLE 8469115
JOURNAL 4 (bases 1 to 24595)
MEDLINE Guevara-Garcia,A., Mosqueda-Cano,G., Arguello-Astorga,G.,
PUBMED Simpson,J. and Herrera-Betrelle,L.
REFERENCE tisbue-specific and wound-inducible pattern of expression of the
AUTHORS mannopine synthase promoter is determined by the interaction
TITLE between positive and negative cis-regulatory elements
JOURNAL Plant J. 4 (3), 495-505 (1993)
MEDLINE 94035196
PUBMED 8220492
FEATURES
source location/Qualifiers
1. .24595
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:358"
/plasmid="Ti"
909..23782
/misc_feature /note="T-region"
909..932
/misc_feature /note="direct repeat A"
1337
/conflict /note="C is T in [2]"
/citation=[2]
1661..2179
CDS /note="open reading frame 1"
/codon_start=1
/transl_table=1
/protein_id="CAA25163.1"
/db_xref="GI:39063"
/db_xref="SWISS-PROT:P04028"
/translatability="MYDGPINFIDSSNLQDRRELKLVLIHTEMAYRSSAQSRLAS
ORSMIFIIINTDVPIDPAKDEYVCSKRVACCPDPTIPDILNVSLVYCSFOEMR
RYAQQRPYDGVSDGAVISTVPPYABGITQTRMLQKVMQNTSKETHDLDAVIAL
PNIRFKIOTISHI"
complement(1736..2113)
CDS /note="open reading frame 2"
/codon_start=1
/transl_table=1
/protein_id="CAA25164.1"
/db_xref="GI:39064"
/db_xref="SPTREMBL:044387"
/translatability="WCFPACILIPNPFILPOPHSLFCYSFGIGMDGDCNCAVNAVIES
LNCIPSHFLKRAINVQORYIEYIKGYICWMAARNFARALDHFVFSWNGNISINDEI
YPPALRSYETSLCRTAISILCVN"
1948
/conflict /note="C is T in [2]"
/citation=[2]
2146
/conflict /note="T is C in [2]"
/citation=[2]
complement(2907..3287)
CDS /note="open reading frame 3"
/codon_start=1
/transl_table=1
/protein_id="CAA25165.1"
/db_xref="GI:39065"
/db_xref="SWISS-PROT:P03867"
/conflict /translatability="MNFADTPLASLDLDMWACEEPIFYTGASPOLTEGEVITQNNGLY
LYCKGSLSORIDTHLKKEKEEISTTTIKRAEMKAOQSDLTYYVAIPQSNVYLCVSN
PEKGFURCKNRPLPIYVAHGSMS"
3369..3370
/conflict /note="TC is TCG in [2]"
/citation=[2]
3810..3811
/conflict /note="AA is A in [2]"
/citation=[2]
complement(4060..5463)
CDS /note="open reading frame 4 gene tma"
/codon_start=1
/transl_table=1
/protein_id="CAA25166.1"
/db_xref="GI:39066"
/db_xref="SWISS-PROT:P03868"
/translatability="MVAITSLSQSLLEHLRKQVSCHELVETLJARCEAKSJNALAT
DMGLRSASAKIDIRHGNAGVLCGILPCFKANATGVFPPTSATPALINLPKIPSR
AERLFPAGALPGASGMHESLPGITSNNTATGAVRRPMNDLIPGSSGGVAASVSR
IMIGITGTPNGASVRLPALCGVGRPTLGRVPGDRITIPVSPTRDPTGIIACVADV
VILDRISGTPERIPPVPLKGRIGLPTTYFDLADADVALAETIRILANKGVTFV
EAVIPHLDELNKGASFVALYEPHALKQYLDVFTVVSFSDVIKQIRSPDVANINA
OIDGHQISAEVELARHSRPRLOATYRVFKLRLDALFPTAPLVARPIGQSSVI
HNGTMIDTEFKIYRVNVPSSNAGLPGLSIPVCLTPDLPGMBIDGLASDQRLAIG
GALEBAIGFRYFAGLPN"
5561
/conflict /note="T is C in [2]"
/citation=[2]
5809..8076
CDS /note="open reading frame 5 gene tma"
/codon_start=1
/transl_table=1
/protein_id="CAA25167.1"
/db_xref="GI:39067"
/db_xref="SPTREMBL:044388"
/translatability="MSASPLDNOCDHLPTKAVDLMVDAKELDRRVSDAFIERENS
RGRRIQISTECGAGLAKRLADGRPEISAGKVVVLSAYIYIGKILGRILIESKM
ARATVSGALVAIDLAPRCMDPSEBOLIOALFVLSGRCAPIDSHFVAISIKYRPGT
LPMPLVENGTMKCVTPTTTBGAVPFDVAVAGRNMLKGSAGSPFTDILTYVRRPF
DOCSDSGRIGFPEDVPRKRVAVTIGAGISGLVAVANLHAGVDVTTIYESDVGRLU
WSHAFRDADSVVAEMGAMRFPAPAFCLFFLEBYGLSMRPFNPQTVDTYLVYQGO
YMKAKOQLPPKLFHRYVNGMRAFLKXGFERDVLASPAVITOLKSGDIRMAHDSGO
IMLNRPRESFSGSIEIRIFLGTHPFGGETMSFPHDVDFKLMGIGSGFGFVPSGFI
ELIRLVINGVEENOMWCPEGISLPRIASEVNAGVVSORICHOVYRAIOKRTKIX
IRKXGISSELYDQVYVYTSGLANTQLRHCLTCQNTINIQAPNQAQVDSHMGSKLEPM
TERKFWLDHILBSCLMDGILAKAVICLDBSQPNRKGVLVITMEDSHKLLAAPD
KERKCLFLDLDAISRSPAPAOHLFPAACADYDQVNIQHDWLITDNAGAGAFKLNRRGDE
YSEELFFOALDNTADTNGVYLAGSCSFTGSGWEAGAIOTACNAVCAIHNCGILAKGN
PLEHSWKRNVYRTRN"
complement(6259..6579)
CDS /note="open reading frame 6"
/codon_start=1
/transl_table=1
/protein_id="CAA25168.1"
/db_xref="GI:39068"
/db_xref="SPTREMBL:044389"
/translatability="MKQFVCHHSSGNASANDRRFRILNIRKKADESTICTLYKXSV
VWEQVDCMGTCKTLOHOVSTISYHVKHGPFGYGRPGNAPFRALILVORHOGSAG
SLRD"
6894..6895
/conflict /note="TC is CT in [2]"
/citation=[2]
7597
/conflict /note="T is C in [2]"
/citation=[2]
7778..8113
CDS /note="open reading frame 7"
/codon_start=1
/transl_table=1
/protein_id="CAA25169.1"
/db_xref="GI:39069"
/db_xref="SPTREMBL:044390"
/translatability="MLFNMIGLQTRWGSLSNSTGVVRIFILNFSFGKHWTRLMIPEF
```

```

TWRAVAVPSQVDNRVLFRRPRVPSVOLSTIVEAFMORALISNTLGRDITTALEISLM
ILQVLQVL"
7961
/nc="T is A in [2]"
/citation=[2]
7964
/nc="A is G in [2]"
/citation=[2]
7968. :7969
/nc="CG is GC in [2]"
/citation=[2]
8387
/nc="A is G in [2]"
/citation=[2]
8771. :9493
/nc="open reading frame 8 gene tmr"
CDS
/codon_start=1
/transl_table=1
/protein_id="CAA25170.1"
/db_xref="GI:39070".
/db_xref="SWISS-PROT:P03869"
/translation="MDLHLFPGPTCTKTTTALAALQOTGLPULSDRVCCPOLSTG
SGRPYVELNGITRLYLDDRLVLEGIIIAAQAHRRLIEVYNHEANGSLILEGISTL
LNCMAKNSVSADEFRMHIIIRKLPDQETPKAAKARVLMHPAAGSIIIOELVYLMN
EPLRPLIKEDIGRYAMLPASQNIADWLQLDANMEGKLINGIAQSEYFIHARQOE
OKPQVVAALFDFEGHPFGMY"
9564
/nc="T is C in [2]"
/citation=[2]
9926. :9927
/nc="T is T in [2]"
/citation=[2]
9996. :10571
/nc="open reading frame 9"
/codon_start=1
/transl_table=1
/protein_id="CAA25171.1"
/db_xref="GI:39071"
/db_xref="SWISS-PROT:P04030"
/translation="MDRMSARQGMVPCISHGKDOLQGEISELSKVYREKQTDLH
TKSGDIIINPGSEFLYILDKENYRLCQRNVLVSNASDGLLATLLEPYSDGYTFRQVR
Query Match 98.4%; Score 2230.6; DB 1; Length 24595;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2258; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
QY 1 ATGTGACCTTCACCTCTCTTGATMACAGTGCATCATCTCCCAACCAAAATGTGGAT 60
DB 5809 ATGTGACCTTCACCTCTCTTGATMACAGTGCATCATCTCCCAACCAAAATGTGGAT 5868
QY 61 CTGACATGTGCTAAGGGGAGTGAATTGGAACCGCGGGTTTCCGATGCTTTTGAA 120
DB 5869 CTGACATGTGCTAAGGGGAGTGAATTGGAACCGCGGGTTTCCGATGCTTTTGAA 5928
QY 121 CGAAGACCTTCTAGGGAGAGGATTAACAAGCTCAACGAGTGCAGCGGCTGTTA 180
DB 5929 CGAAGACCTTCTAGGGAGAGGATTAACAAGCTCAACGAGTGCAGCGGCTGTTA 5988
QY 181 GCTTGCAAAAGGTGCGCATGTGCTTCCCGAGATCTCAGCTGTGAAAAGTGA 240
DB 5989 GCTTGCAAAAGGTGCGCATGTGCTTCCCGAGATCTCAGCTGTGAAAAGTGA 6048
QY 241 GTTCTCTCCGCTTATATCTATTTGGCAAGAAATTCGGGGCGGATACTTGAATGAAA 300
DB 6049 GTTCTCTCCGCTTATATCTATTTGGCAAGAAATTCGGGGCGGATACTTGAATGAAA 6108
QY 301 CTTGGGCGGGGCAACAGTGAAGTGTCTGCTGSCATCGACTTGACACCATTTGGCATG 360
DB 6109 CTTGGGCGGGGCAACAGTGAAGTGTCTGCTGSCATCGACTTGACACCATTTGGCATG 6168
QY 361 GATTTCCTCCGAAGACAACATAATCCAGCCCTGTTTTTGTGAGCGGTAAAGATGTGCA 420
DB 6169 GATTTCCTCCGAAGACAACATAATCCAGCCCTGTTTTTGTGAGCGGTAAAGATGTGCA 6228
```

```

CCGATGATCTTAGTCAATTTGNGGCCATTTCAATCTCTAAGACTGCGGGCTTTGCAACC 480
DB 421 CCGATGATCTTAGTCAATTTGNGGCCATTTCAATCTCTAAGACTGCGGGCTTTGCAACC 6228
QY 481 CTGCAATGCGCGCTGTAAGAAATGCGACGATGAATGCGTTACCGGGTTTACATTAACC 540
DB 6289 CTGCAATGCGCGCTGTAAGAAATGCGACGATGAATGCGTTACCGGGTTTACATTAACC 6348
QY 541 CTTGAAGGGGCGGCTGCCATTTGACATGTAAGCTTATAGTTGAAACCTGATGCTGAAGGTT 600
DB 6349 CTTGAAGGGGCGGCTGCCATTTGACATGTAAGCTTATAGTTGAAACCTGATGCTGAAGGTT 6408
QY 601 TCGGCAAGGTTCTTTCCAAACAATCGACTTCTCTAAGACTGCAAGCCGTTTTTGAACCA 660
DB 6409 TCGGCAAGGTTCTTTCCAAACAATCGACTTCTCTAAGACTGCAAGCCGTTTTTGAACCA 6468
QY 661 TGTTCGATAGTGAAGGATCGGCTTCTTCCGAGAGATGTTCTAAGCCGAAAGTGGCG 720
DB 6469 TGTTCGATAGTGAAGGATCGGCTTCTTCCGAGAGATGTTCTAAGCCGAAAGTGGCG 6528
QY 721 GTCATTTGGCGCTGCAATTTCCGCACTGCTGCTGGCAAGCAACTGCTTCAATGCTGGGGTA 780
DB 6529 GTCATTTGGCGCTGCAATTTCCGCACTGCTGCTGGCAAGCAACTGCTTCAATGCTGGGGTA 6588
QY 781 GAGGATGTTACATATATGAAGCAAGTGAATCGTGTGGAGGCAAGCTTTGTCATAGCT 840
DB 6589 GAGGATGTTACATATATGAAGCAAGTGAATCGTGTGGAGGCAAGCTTTGTCATAGCT 6648
QY 841 TTCAGGGAACGCTCTAGTGTGCTGCGCAAAATGGGAGCGATGCGATTTCTCTGCTGCA 900
DB 6649 TTCAGGGAACGCTCTAGTGTGCTGCGCAAAATGGGAGCGATGCGATTTCTCTGCTGCA 6708
QY 901 TTTGCTGTTTTTTCTTCTCGAGCGTTACGAGCTGTCTTGATGAGGCGTTCCCAAT 960
DB 6709 TTTGCTGTTTTTTCTTCTCGAGCGTTACGAGCTGTCTTGATGAGGCGTTCCCAAT 6768
QY 961 CCGGGAACGCTGCACTTAATGCTGTCAACAGGGGTCACAAACATATGTAAGCCGGG 1020
DB 6769 CCGGGAACGCTGCACTTAATGCTGTCAACAGGGGTCACAAACATATGTAAGCCGGG 6628
QY 1021 CAGCTCCACCGAAGCTGTTCCATCGCGTTTCAACGAGTTGGCGTCTTCTTGAAGAC 1080
DB 6829 CAGCTCCACCGAAGCTGTTCCATCGCGTTTCAACGAGTTGGCGTCTTCTTGAAGAC 6888
QY 1081 GGTTCATGAGCGAATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 6889 GGTTCATGAGCGAATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6948
QY 1141 TCAGACACATTAAGTGGGCTCATGACTCTGCGCAATTTGGCTGAACGCTTTGGGAGG 1200
DB 6949 TCAGACACATTAAGTGGGCTCATGACTCTGCGCAATTTGGCTGAACGCTTTGGGAGG 7008
QY 1201 GAGTCTTCTCTCAGAGGATAGAGAGATCTTCTGAGGACACACTCTCTGAGTGT -AA 1259
DB 7009 GAGTCTTCTCTCAGAGGATAGAGAGATCTTCTGAGGACACACTCTCTGAGTGTGAA 7068
QY 1260 ACATGAGATTTTCTCATGATTTGGGACTTAATCAAGCTTAATGAGGATAGAGATCTGCGGG 1319
DB 7069 ACATGAGATTTTCTCATGATTTGGGACTTAATCAAGCTTAATGAGGATAGAGATCTGCGG -GG 7127
QY 1320 GTTGTGCTCAAGTTTGAAGCGGGTTTATTTGAGATCTCTGCTGCTGCTCATCAAGGATA 1379
DB 7128 GTTGTGCTCAAGTTTGAAGCGGGTTTATTTGAGATCTCTGCTGCTGCTCATCAAGGATA 7187
QY 1380 TGAAGAAATACAGCGGATGCGCTGAGAGAACTCTGAGAACTTCCAGCTGAGATGCGATC 1439
DB 7188 TGAAGAAATACAGCGGATGCGCTGAGAGAACTCTGAGAACTTCCAGCTGAGATGCGATC 7247
QY 1440 TGAAGTGTAAACGATGTGTCTGTGAGCAAGCGCATATGCGCATGTTCAAGTCAAGGCGAT 1499
DB 7248 TGAAGTGTAAACGATGTGTCTGTGAGCAAGCGCATATGCGCATGTTCAAGTCAAGGCGAT 7307
```

```

QY 1500 TCAGAGGAAAGACAAATAATTAAGATTAGGCTTAAGCGCGGATATCTGAATCTTATGA 1559
DB 7308 TCAGAGGAAAGACAAATAATTAAGATTAGGCTTAAGCGCGGATATCTGAATCTTATGA 7367
QY 1560 TAAAGTGTGTGTCATCTGTGACTCGCAAAATATCAACTCGAGGATTTGCTGACATGCGA 1619
DB 7368 TAAAGTGTGTGTCATCTGTGACTCGCAAAATATCAACTCGAGGATTTGCTGACATGCGA 7427
QY 1620 TACCAATATTTTATGAGGACCAAGTGAACCAAGCGGTTGATTAACGCCATATGACAGATC 1679
DB 7428 TACCAATATTTTATGAGGACCAAGTGAACCAAGCGGTTGATTAACGCCATATGACAGATC 7487
QY 1680 GTCAAAAATCTTCTCTGATGACTGAACGAAAAATTCGTGTTAGACCATATCCCTCCGCTTG 1739
DB 7488 GTCAAAAATCTTCTCTGATGACTGAACGAAAAATTCGTGTTAGACCATATCCCTCCGCTTG 7547
QY 1740 TGTCTCATGAGACGGGATCGCAAAAGCATGTATTCCTGAGCTATGATGACAGATCC 1799
DB 7548 TGTCTCATGAGACGGGATCGCAAAAGCATGTATTCCTGAGCTATGATGACAGATCC 7607
QY 1800 GAATGTGAAGGTCTAGTGTCTCATGATTAACATGAGGAGACGACTCCCAAGCTGTT 1859
DB 7608 GAATGTGAAGGTCTAGTGTCTCATGATTAACATGAGGAGACGACTCCCAAGCTGTT 7667
QY 1860 GGGCGTCCCGACAAAAAAGAGCATATGCTGCTGCGGAGCAATTTGAGATCTTT 1919
DB 7668 GGGCGTCCCGACAAAAAAGAGCATATGCTGCTGCGGAGCAATTTGAGATCTTT 7727
QY 1920 CCGCGCTTTGCGGACGACCTATTTCTGCTGCGGATTAACGCAAAATGTTATTC 1979
DB 7728 CCGCGCTTTGCGGACGACCTATTTCTGCTGCGGATTAACGCAAAATGTTATTC 7787
QY 1980 ACATGATGTGCTTACACAGCAAAATGCGGAGGAGCTTTCAAACTCAACCGGCTGTGA 2039
DB 7788 ACATGATGTGCTTACACAGCAAAATGCGGAGGAGCTTTCAAACTCAACCGGCTGTGA 7847
QY 2040 GGATTTTATTTCTGAAGAACTTTCTTCAAGCACTGAGCAACGCGCTATGATACCGAGT 2099
DB 7848 GGATTTTATTTCTGAAGAACTTTCTTCAAGCACTGAGCAACGCGCTATGATACCGAGT 7907
QY 2100 TTAATTGCGGCGGTGCAATTTGCTTCAACAGTGAATGAGTGGAGGCTGCTATTCAGAC 2159
DB 7908 TTAATTGCGGCGGTGCAATTTGCTTCAACAGTGAATGAGTGGAGGCTGCTATTCAGAC 7967
QY 2160 CGCGTGAAGCGGCTCTGTGCAATTAATCCCAATTTGAGAGGCTTTTGGCAAAAGGCGA 2219
DB 7968 CGCGTGAAGCGGCTCTGTGCAATTAATCCCAATTTGAGAGGCTTTTGGCAAAAGGCGA 8027
QY 2220 TCCCTCGAACAATCTTGAAGAGATTAATTAACCGGAGTAGAATATG 2268
DB 8028 TCCCTCGAACAATCTTGAAGAGATTAATTAACCGGAGTAGAATATG 8076

RESULT 2
BD016312 24595 bp DNA linear PAT 27-AUG-2002
LOCUS Method of promoting plant transcription by using octopline T-DNA
DEFINITION
ACCESSION BD016312
VERSION BD016312.1 GI:22557450
KEYWORDS JP 2001190289-A/1.
SOURCE Agrobacterium tumefaciens (Rhizobium radiobacter)
ORGANISM Agrobacterium tumefaciens
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
REFERENCE 1 (bases 1 to 24595)
AUTHORS Barker R.F. and Kemp J.D.
TITLES Method of promoting plant transcription by using octopline T-DNA
JOURNAL Patent: JP 2001190289-A 1 17-JUL-2001;
MYCOGEN PLANT SCIENCE INC
OS Agrobacterium tumefaciens
PN JP 2001190289-A/1
PD 17-JUL-2001

```

```

PF 22-NOV-2000 JP 2000356816
PR 18-NOV-1983 US 553786
PI RICHARD F BARKER, JOHN D KEMP
PC C12N15/09,A01H5/00,C12N5/10,C12N5/00,C12N5/00 CC Method of
promoting plant transcription by using octopline T-CC
DNA promoter
FH Key 1.24595
FT source location/Qualifiers
location/Qualifiers
1.24595 /organism='Agrobacterium tumefaciens'
/organism='Agrobacterium tumefaciens'
/mol type='genomic DNA'
/db xref='taxon:358'
BASE COUNT 6534 a 5509 c 5793 g 6759 t
ORIGIN
Query Match 98.4%; Score 2230.6; DB 6; Length 24595;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2258; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
QY 1 ATGTCAAGCTTCACTCTCTTGTATTAACAGTGCATCATCTCCCAACCAAAATGTTGAT 60
DB 5809 ATGTCAAGCTTCACTCTCTTGTATTAACAGTGCATCATCTCCCAACCAAAATGTTGAT 5868
QY 61 CTGCAATGCTGATTAAGCGGAGTGAATGGAACGCGAGGTTCCGATGCTTTAGAA 120
DB 5869 CTGCAATGCTGATTAAGCGGAGTGAATGGAACGCGAGGTTCCGATGCTTTAGAA 5928
QY 121 CGAAGACTTCTAGGAGGAGAGATTACTCAAAAGCTCCACGAGTGCAGCGTGGGTTA 180
DB 5929 CGAAGACTTCTAGGAGGAGAGATTACTCAAAAGCTCCACGAGTGCAGCGTGGGTTA 5988
QY 181 GCTTGAAGAGGCTGCGCATGCTGCTTCCCGAATCTCACTGTGGAAGGTGCA 240
DB 5989 GCTTGAAGAGGCTGCGCATGCTGCTTCCCGAATCTCACTGTGGAAGGTGCA 6048
QY 241 GTTCTCCGCTTATATATTAATGGAAGAAATCTGGGAGGAGTACTGAATGAAA 300
DB 6049 GTTCTCCGCTTATATATTAATGGAAGAAATCTGGGAGGAGTACTGAATGAAA 6108
QY 301 CTTTGGGCGGCGCAACAGTGAAGTGTCTGTCATGCACTGATGCAACCATTTGATG 360
DB 6109 CTTTGGGCGGCGCAACAGTGAAGTGTCTGTCATGCACTGATGCAACCATTTGATG 6168
QY 361 GATTTCTCGAAGCACTAATCAAAAGCTGTTTGTGTCAGCGGTAAAGATGCA 420
DB 6169 GATTTCTCGAAGCACTAATCAAAAGCTGTTTGTGTCAGCGGTAAAGATGCA 6228
QY 421 CGATTTGATCTTATGATTTTGTGAGCAATTTCAATCTTAAGATGCGCGCTTTCGAA 480
DB 6229 CGATTTGATCTTATGATTTTGTGAGCAATTTCAATCTTAAGATGCGCGCTTTCGAA 6288
QY 481 CTGCAATGCGCGCTGTAAGAGATGCGAGATGAATGCGGTTACCGGTTTACATACC 540
DB 6289 CTGCAATGCGCGCTGTAAGAGATGCGAGATGAATGCGGTTACCGGTTTACATACC 6348
QY 541 CTTGAAGGCGCGCTGCAATTTGACATGTTAGCTTATGTTGAAAACCTGATGCTGAAGGT 600
DB 6349 CTTGAAGGCGCGCTGCAATTTGACATGTTAGCTTATGTTGAAAACCTGATGCTGAAGGT 6408
QY 601 TCGGCAAGGTTCTTCCCAACATCGACTGCTAGAGCTGAGACCGGTTTGTGACCA 660
DB 6409 TCGGCAAGGTTCTTCCCAACATCGACTGCTAGAGCTGAGACCGGTTTGTGACCA 6468
QY 661 TGTTCGATAGTGAAGCATGCGCTTCTTCCGAGAGATGTTCTTAAGCGGAAGTGGCG 720
DB 6469 TGTTCGATAGTGAAGCATGCGCTTCTTCCGAGAGATGTTCTTAAGCGGAAGTGGCG 6528
QY 721 GTCAATGCGCGCTGCAATTTCCGAGCTGTTGTTGCAACGAACTGCTTCACTGCGGCTA 780
DB 6529 GTCAATGCGCGCTGCAATTTCCGAGCTGTTGTTGCAACGAACTGCTTCACTGCGGCTA 6588

```

OY	781	GACGATGTACAAATAATATGAAGCAAGTATCGTGTGGAGGCAAGCTTTGGTCAATGCT	840
Db	6589	GACGATGTACAAATAATATGAAGCAAGTATCGTGTGGAGGCAAGCTTTGGTCAATGCT	6644
OY	841	TTTCAGGAGCGCTTCTAGTGTGTGGCGAAATGGGGGCGATCGAATTTCTCTGTGTGA	900
Db	6649	TTTCAGGAGCGCTTCTAGTGTGTGGCGAAATGGGGGCGATCGAATTTCTCTGTGTGA	6700
OY	901	TTCTGCTTTTCTTTCTTCTCTCGAGCGTTACGGCCTGTCTTTCGATGAGGCCGTTCCCAAT	960
Db	6709	TTCTGCTTTTCTTTCTTCTCTCGAGCGTTACGGCCTGTCTTTCGATGAGGCCGTTCCCAAT	6766
OY	961	CCCGGACAGTGGACACTACTTGGTGTACCAAGGGTCCAAATACATGTGAAAGCCGGG	1022
Db	6769	CCCGGACAGTGGACACTACTTGGTGTACCAAGGGTCCAAATACATGTGAAAGCCGGG	6822
OY	1021	CAGCTGCCACCGAAGCTGTTCCATCGCGTTTACCAACGGTTGGCGTGTCTTTGAAGAC	1080
Db	6829	CAGCTGCCACCGAAGCTGTTCCATCGCGTTTACCAACGGTTGGCGTGTCTTTGAAGAC	6888
OY	1081	GCTTTTCATGAGCGAGATATGTGTGGCTTGCGCTTGCTATTACTAGGCCCTTTGAAA	1144
Db	6889	GCTTTTCATGAGCGAGATATGTGTGGCTTGCGCTTGCTATTACTAGGCCCTTTGAAA	6946
OY	1141	TCAGAGACATTAGGTGGGCTCATGACTCTCTGGCAAATTTGGCTGAACGTTTCGGGAGG	1200
Db	6949	TCAGAGACATTAGGTGGGCTCATGACTCTCTGGCAAATTTGGCTGAACGTTTCGGGAGG	7008
OY	1201	GAGTCCCTTCTCTTCAGGGGATAGAGAGATCTTTCGGGACACATCTCTGGTGTGT-AA	1255
Db	7009	GAGTCCCTTCTCTTCAGGGGATAGAGAGATCTTTCGGGACACATCTCTGGTGTGTAA	7066
OY	1260	ACATGAGTTTCTCTCATGATTTGGAGCCTATTCAAAGCTAATGGAAATAGATCTTGGCGGG	1319
Db	7069	ACATGAGTTTCTCTCATGATTTGGAGCCTATTCAAAGCTAATGGAAATAGATCTTGGCGGG	7122
OY	1320	GTTTGGTCCAGTTTGTGAAGGGGGTTTATTGAGATCTCTCGGCTTGGTCAATCAGGATA	1379
Db	7128	GTTTGGTCCAGTTTGTGAAGGGGGTTTATTGAGATCTCTCGGCTTGGTCAATCAGGATA	7187
OY	1380	TGAAGAAATCAGCGGATGTGCGCTTGAAGGAAATCTCAGAACCTTCCAGTCCGATCGCATC	1439
Db	7188	TGAAGAAATCAGCGGATGTGCGCTTGAAGGAAATCTCAGAACCTTCCAGTCCGATCGCATC	7247
OY	1440	TGAATGTGTAAACGGTGTCTGTGAGCCAGCGCATATGTCATGTTCAAGTCAGGGCAT	1499
Db	7248	TGAATGTGTAAACGGTGTCTGTGAGCCAGCGCATATGTCATGTTCAAGTCAGGGCAT	7307
OY	1500	TCAGAGGAAAGACAAAAATTAAGATTAAGGCTTAAAGCCGGGATATCTGAATCTTTATGA	1559
Db	7308	TCAGAGGAAAGACAAAAATTAAGATTAAGGCTTAAAGCCGGGATATCTGAATCTTTATGA	7367
OY	1560	TAAAGTGTGGTGCACATCTGGACCTGGCAAAATATCCAACTCAGGATTCGCTGACATGGA	1619
Db	7368	TAAAGTGTGGTGCACATCTGGACCTGGCAAAATATCCAACTCAGGATTCGCTGACATGGA	7427
OY	1620	TACCAATATTTTTCAGGCACCAAGTGAACCAAGCGGTGTATTAACAGCCATATGACAGATC	1679
Db	7428	TACCAATATTTTTCAGGCACCAAGTGAACCAAGCGGTGTATTAACAGCCATATGACAGATC	7487
OY	1680	GTCAAAATCTTCTCTGATGACTGAACGAAAAATCTGTGTAGAACCATATCTCCGCTTGG	1739
Db	7488	GTCAAAATCTTCTCTGATGACTGAACGAAAAATCTGTGTGTAGAACCATATCTCCGCTTGG	7547
OY	1740	TGTCCTCATGAGACGGGATCGCAAAAGCGGTATTTGCTGGACATATAGTGGCAGGATCC	1799
Db	7548	TGTCCTCATGAGACGGGATCGCAAAAGCGGTATTTGCTGGACATATAGTGGCAGGATCC	7607
OY	1800	GAATGTTAAGGTCTAGTGTCTCATCAGTTATATATGAGGAGACGATCTCCCAAGCTGTT	1859
Db	7608	GAATGTTAAGGTCTAGTGTCTCATCAGTTATATATGAGGAGACGATCTCCCAAGCTGTT	7667
OY	1860	GGCGGTCCCGACAAAAAAGCGCATATATGCTGTGCTGGGAGCGCAATTTTCAGATCTTT	1919

Db	LOCUS	DEFINITION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	JOURNAL	COMMENT
Db	7668	GGCGGTCCCGACAAAAGGCATTAATGCTGCGGGAGCGCAATTCGAGATCTTT							
Qy	1920	CCCGGCGTTTGCCGAGCACCATTATTCGTGCGTGGCGTGAATTACGACCCAAAATGTATTCA	1979						
Db	7728	CCCGGCGTTTGCCGAGCACCATTATTCGTGCGTGGCGTGAATTACGACCCAAAATGTATTCA	7787						
Qy	1980	ACATGATTGGCTTACAGACGAGAAATGCCGGGGAGCTTTCAAACTCAACCGGCGGTGTA	2039						
Db	7788	ACATGATTGGCTTACAGACGAGAAATGCCGGGGAGCTTTCAAACTCAACCGGCGGTGTA	7847						
Qy	2040	GGATTTTATTTCGAGAACTTTTCTTTCAAGCACTGAGACCGGCTPATGATACCGGAGT	2099						
Db	7848	GGATTTTATTTCGAGAACTTTTCTTTCAAGCACTGAGACCGGCTPATGATACCGGAGT	7907						
Qy	2100	TTACTTGGCGGGTTGACGTGTTCTTCAAGGNGATGGGGTGGAGGTCATATTGACAG	2159						
Db	7908	TTACTTGGCGGGTTGACGTGTTCTTCAAGGNGATGGGGTGGAGGTCATATTGACAG	7967						
Qy	2160	CGCGTGTACGCGGTCTGTGCAATTATCCCAATTGTGAGGCAATTTGGCAAGGCGCA	2219						
Db	7968	CGCGTGTACGCGGTCTGTGCAATTATCCCAATTGTGAGGCAATTTGGCAAGGCGCA	8027						
Qy	2220	TCTCTTGAACTCTTTGGAAGATATTAATACCCGAGTAGAAATTAG	2268						
Db	8028	TCTCTTGAACTCTTTGGAAGATATTAATACCCGAGTAGAAATTAG	8076						
RESULT 3									
E00404	LOCUS	E00404	24595 bp	DNA	linear	PAT 29-SEP-1997			
E00404	DEFINITION	T1 plasmid T-DNA region.							
E00404	VERSION	E00404.1 GI:2168687							
E00404	KEYWORDS	JP 1985156333-A/1.							
E00404	ORGANISM	Agrobacterium tumefaciens (Rhizobium radiobacter)							
E00404	REFERENCE	Agrobacterium tumefaciens (Rhizobium radiobacter); Rhizobiales; Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium. 1 (bases 1 to 24595)							
E00404	AUTHORS	Garll,E.D., Denisu,D.S. and RichiYaado,E.B.							
E00404	JOURNAL	SELECTION USING OPINE SYNTHASE GENE							
E00404	COMMENT	Patent: JP 1985156333-A 1 16-AUG-1985; ACURIJIENTEIKUSU RES ASSOC LTD OS Agrobacterium tumefaciens PN JP 1985156333-A/1							
E00404	PD	16-AUG-1985							
E00404	PF	14-SEP-1984 JP 1984193841							
E00404	PI	14-SEP-1983 US 83 532280							
E00404	PC	GARLI EE DARU, DENISU DABURIYU SATSUTON, RICHIIYAADO EBU BEIKAA							
E00404	CC	A01H1/00,C12N5/00,C12N15/00,(C12N5/00,C12N1:91); CC							
E00404	CC	strandedness: Double;							
E00404	CC	topology: Linear;							
E00404	CC	*source: strain=15955;							
E00404	FM	Key							
E00404	FM	Location/Qualifiers							
E00404	FM	repeat_unit	909..1032						
E00404	FM	/note='ROTL(A)'							
E00404	FM	repeat_unit	14060..14083						
E00404	FM	/note='ROTL(B)'							
E00404	FM	repeat_unit	15900..15923						
E00404	FM	/note='ROTL(C)'							
E00404	FM	repeat_unit	23759..23783						
E00404	FM	/note='ROTL(D)'							
E00404	FM	complement(12555..13632)							
E00404	FM	/product='ocropine synthase'							
E00404	FM	complement(19147..20128)							
E00404	FM	/product='agropine synthase'							
E00404	FM	20611..21900							
E00404	FM	/product='agropine synthase'							
E00404	FM	complement(22482..23384)							
E00404	FM	/product='mannopine synthase'.							

FEATURES
source Location/Qualifiers
1. 24595
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:358"
BASE COUNT 6534 a 5509 c 5793 g 6759 t
ORIGIN
Query Match 98.4%; Score 2230.6; DB 6; Length 24595;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2258; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
QY 1 ATGTCACTTCACTCTCTTGATTAACAGTGGCATCTCCCAACCAAAATGTGAT 60
Db 5809 ATGTCACTTCACTCTCTTGATTAACAGTGGCATCTCCCAACCAAAATGTGAT 5868
QY 61 CTGACAAATGTCATTAAGCGGATGAATTGACCGCAGGGTTTCCATGCTTTAGAA 120
Db 5869 CTGACAAATGTCATTAAGCGGATGAATTGACCGCAGGGTTTCCATGCTTTAGAA 5928
QY 121 CGAAGAGCTTCTAGGGGAGAGAGATTACCAAGCTCCACCGAGTGCAGCGCTGGTTA 180
Db 5929 CGAAGAGCTTCTAGGGGAGAGAGATTACCAAGCTCCACCGAGTGCAGCGCTGGTTA 5988
QY 181 GCTTGAAGAGCTGCGCATGTGCTCCCGAGATCTCAGCTGTGGAAGTAGCA 240
Db 5989 GCTTGAAGAGCTGCGCATGTGCTCCCGAGATCTCAGCTGTGGAAGTAGCA 6048
QY 241 GTTCTCTCGCTTATATCTATATTGGCAAGAAATCTGGGCGGATCTTGAATGAAA 300
Db 6049 GTTCTCTCGCTTATATCTATATTGGCAAGAAATCTGGGCGGATCTTGAATGAAA 6108
QY 301 CCTTGGCGCGGGGCAAGAGATGCTCTGTTGCAATGCACTTGGACACATTTGCATG 360
Db 6109 CCTTGGCGCGGGGCAAGAGATGCTCTGTTGCAATGCACTTGGACACATTTGCATG 6168
QY 361 GATTTCCTCGAAGACAACTAATCCAAGCCCTGTTTTGCTGACGGTAAAGATGCA 420
Db 6169 GATTTCCTCGAAGACAACTAATCCAAGCCCTGTTTTGCTGACGGTAAAGATGCA 6228
QY 421 CCGATTGATCTTAGTCAATTTGCTGGCCATTCAATCTTAAGACTGCGGCTTTGCAAC 480
Db 6229 CCGATTGATCTTAGTCAATTTGCTGGCCATTCAATCTTAAGACTGCGGCTTTGCAAC 6288
QY 481 CTGCGCAATGCGGCTTACAGAGATGAGACAGTAAAGCGTTACCGGTTTACATTAAC 540
Db 6289 CTGCGCAATGCGGCTTACAGAGATGAGACAGTAAAGCGTTACCGGTTTACATTAAC 6348
QY 541 CTGAAGAGGGCGTGCATTGACATGTAGCTTAATGTGGAACCTGATGCTGAAGGT 600
Db 6349 CTGAAGAGGGCGTGCATTGACATGTAGCTTAATGTGGAACCTGATGCTGAAGGT 6408
QY 601 TCGGAGAGTTCTTTCCAAACATTCGACTTACGACTGACAGCCGTTTTTTGACCA 660
Db 6409 TCGGAGAGTTCTTTCCAAACATTCGACTTACGACTGACAGCCGTTTTTTGACCA 6468
QY 661 TGTTCGATAGTGAAGGATCGGCTTTCCGAGAGATTTCTTAAGCCGGAAGTGGCG 720
Db 6469 TGTTCGATAGTGAAGGATCGGCTTTCCGAGAGATTTCTTAAGCCGGAAGTGGCG 6528
QY 721 GTATTGGCGCTGGCATTTCCGACTGATGAGCAACGACTTCACTGCTGGGTTA 780
Db 6529 GTATTGGCGCTGGCATTTCCGACTGATGAGCAACGACTTCACTGCTGGGTTA 6588
QY 781 GACGATTTAACAATATGAAGCAAGTATGTTGTTGAGCAAGCTTTGATCAATGCT 840
Db 6589 GACGATTTAACAATATGAAGCAAGTATGTTGTTGAGCAAGCTTTGATCAATGCT 6648
QY 841 TTCAAGGAGCGCTCTATGTCTGTGGCGGAATGGGGCGATGCAATTTCTCTGCTGCA 900
Db 6649 TTCAAGGAGCGCTCTATGTCTGTGGCGGAATGGGGCGATGCAATTTCTCTGCTGCA 6708
QY 901 TTCTGCTGTTTTTCTTCTCTGAGCGTTACGCGCTGTCTGATGAGCGCTTCCAAAT 960

Db 6709 TTCTGCTGTTTTTCTTCTCTGAGCGTTACGCGCTGTCTGATGAGCGCTTCCAAAT 6768
QY 961 CCCGGCAGGTGACACTTACTTGTCTACCAAGGCGTCCAAATACATGTGAAGCGGG 1020
Db 6769 CCCGGCAGGTGACACTTACTTGTCTACCAAGGCGTCCAAATACATGTGAAGCGGG 6828
QY 1021 CAGCTGCCACCGAAGCTGTTCCATCGCGTTTAAACACGATGTCGCTGTTTGAAGAC 1080
Db 6829 CAGCTGCCACCGAAGCTGTTCCATCGCGTTTAAACACGATGTCGCTGTTTGAAGAC 6888
QY 1081 GGTTCATGACGAGATATTTGTGTTGGCTTGCTGTGCTATTACTAGGCTTTGAAA 1140
Db 6889 GGTTCATGACGAGATATTTGTGTTGGCTTGCTGTGCTATTACTAGGCTTTGAAA 6948
QY 1141 TCAGGACATATTAGGGGCTCATGATCTGCGCAAAATTTGGCTGAACGTTTGGGAG 1200
Db 6949 TCAGGACATATTAGGGGCTCATGATCTGCGCAAAATTTGGCTGAACGTTTGGGAG 7008
QY 1201 GAGTCTTCTCTTCAAGGATAGAGAGATCTTCTGGGACACATCTCTGTTGTT -AA 1259
Db 7009 GAGTCTTCTCTTCAAGGATAGAGAGATCTTCTGGGACACATCTCTGTTGTTGAA 7068
QY 1260 ACATGAGTTTTCTCATGATTTGGAACCTTATCAAGCTAATGGAATAGATCTGGCG 1319
Db 7069 ACATGAGTTTTCTCATGATTTGGAACCTTATCAAGCTAATGGAATAGATCTGGCG 7127
QY 1320 GTTGGTCCAGTTTTTGAAGGGGTTTATGAGATCTCCGCTGGTCAACAGGATA 1379
Db 7128 GTTGGTCCAGTTTTTGAAGGGGTTTATGAGATCTCCGCTGGTCAACAGGATA 7187
QY 1380 TGAAGAAATCAAGCGATGTCCTGAAGAAATCTCAGAACTTCCAGTCCGATC 1439
Db 7188 TGAAGAAATCAAGCGATGTCCTGAAGAAATCTCAGAACTTCCAGTCCGATC 7247
QY 1440 TGAAGTGTAAACGTTGTCTGTGACCGGCAATATGCTCATGTTCAAGTCAAGGCGAT 1499
Db 7248 TGAAGTGTAAACGTTGTCTGTGACCGGCAATATGCTCATGTTCAAGTCAAGGCGAT 7307
QY 1500 TCGAAGGAAAGACAAATAAAGTAAAGCTTAAGAGGGGATATCGAACTTTATGA 1559
Db 7308 TCGAAGGAAAGACAAATAAAGTAAAGCTTAAGAGGGGATATCGAACTTTATGA 7367
QY 1560 TAAAGTGTGTGTCACATCTGGAATCGCAATATCAACTCAGGCAATGCTGATGCGA 1619
Db 7368 TAAAGTGTGTGTCACATCTGGAATCGCAATATCAACTCAGGCAATGCTGATGCGA 7427
QY 1620 TACCAATATTTTCAAGCACAGTGAACCAAGCGTTGATTAACGCCATATGACAGATC 1679
Db 7428 TACCAATATTTTCAAGCACAGTGAACCAAGCGTTGATTAACGCCATATGACAGATC 7487
QY 1680 GTCAAAACCTTCCTGATGATGGAAGAAATTTCTGTTAGACCATATCTCCGCTTG 1739
Db 7488 GTCAAAACCTTCCTGATGATGGAAGAAATTTCTGTTAGACCATATCTCCGCTTG 7547
QY 1740 TGTCTCATGACGCGATGCAAAACAGTATATCTCTGACATATAGTTCGAGATCC 1799
Db 7548 TGTCTCATGACGCGATGCAAAACAGTATATCTCTGACATATAGTTCGAGATCC 7607
QY 1800 GAATGTTAAAGTCTAGTGTCTATCAATTATCATGAGAGAGACCTCCCAAGCTGTT 1859
Db 7608 GAATGTTAAAGTCTAGTGTCTATCAATTATCATGAGAGAGACCTCCCAAGCTGTT 7667
QY 1860 GGGGTCCTCCGCAAAAGAGAGATTATGTCGTCGGGAGGCAATTTGAGATCTT 1919
Db 7668 GGGGTCCTCCGCAAAAGAGAGATTATGTCGTCGGGAGGCAATTTGAGATCTT 7727
QY 1920 CCCGGGTTTCCAGACACTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATCA 1979
Db 7728 CCCGGGTTTCCAGACACTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATCA 7787
QY 1980 ACATGATGCTTACAGACGAGATGCGGGGAGCTTTCAAACTCAACGCGGTGTTGA 2039

Db 7788 ACATGATGGCTTACAGACGAAATGCGGGAGCTTTTAACTCAACCGCGCTGTGA 7847
Qy 2040 GGAATTTTATTCGAAAGACTTTTCTTCAAGCACTGGACACGGCTAATGATACCGGAGT 2099
Db 7848 GGAATTTTATTCGAAAGACTTTTCTTCAAGCACTGGACACGGCTAATGATACCGGAGT 7907
Qy 2100 TTACTTGGCGGGTTCAGTGTTCCTTCAAGGTGATGGGTGAGGGTCTATTTCAGAC 2159
Db 7908 TTACTTGGCGGGTTCAGTGTTCCTTCAAGGTGATGGGTGAGGGTCTATTTCAGAC 7967
Qy 2160 CCGGTGTAAAGCGGTCTGTGCAATTATCCACAAATTTGGAGGCAATTTGGCAAGGCGAA 2219
Db 7968 CCGGTGTAAAGCGGTCTGTGCAATTATCCACAAATTTGGAGGCAATTTGGCAAGGCGAA 8027
Qy 2220 TCCCTCGAAGCACTCTTGAAGATATTAATCTACCGAGTAATAG 2268
Db 8028 TCCCTCGAAGCACTCTTGAAGATATTAATCTACCGAGTAATAG 8076

RESULT 4
E00546 24595 bp DNA linear PAT 29-SEP-1997
LOCUS E00546
DEFINITION DNA fragment of T-DNA of Ti-plasmid, pTI15955.
ACCESSION E00546
VERSION E00546.1 GI:2168825
KEYWORDS JP 1985256383-A/1.
SOURCE Agrobacterium tumefaciens (Rhizobium radiobacter)
ORGANISM Agrobacterium tumefaciens
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
1 (bases 1 to 24595)
RichiYaado, E.B. and Jiyon, D.K.
PROMOTOR OF OCTOPINE T-DNA
Patent: JP 1985256383-A 1 18-DEC-1985;
AGURIJENETIKUSU RES ASSOC LTD
OS Agrobacterium tumefaciens
PN JP 1985256383-A/1
PD 18-DEC-1985
PE 19-NOV-1984 JP 1984244307
PR 18-NOV-1983 US 83 553786
PI RICHIVADO EFU BEIKAA, JIYON DEI KENPU
PC C12N15/00, A01H1/00, C12N1/20, C12N5/00//C07H21/04, (C12N1/20, PC
C12R1/01),
PC (C12N5/00, C12R1/91);
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH
FT CDS >1..<24595
FT /gene='tme'
FT >1..<24595
FT /gene='tme'
FT >1..<24595
FT /gene='tmi'
FT >1..<24595
FT /gene='tmi'
FT CDS >1..<24595
FT /gene='ocs'
FT /product='octopine synthetase'.
FEATURES
source location/Qualifiers
1..24595
/organism='Agrobacterium tumefaciens'
/mol_type='genomic DNA'
/db_xref='taxon:358'

BASE COUNT 6534 a 5509 c 5793 g 6759 t
ORIGIN
Query Match 98.4%; Score 2230.6; DB 6; Length 24595;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2258; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
Qy 1 ATGTGAGCTTCACTCTCTTGTATACAGTGGATCATCTCCCAACCAAAATGTGAT 60
|||||

Db 5809 ATGTGAGCTTCACTCTCTTGTATACAGTGGATCATCTCCCAACCAAAATGTGAT 5868
Qy 61 CTGCAATGTGTGATAGGCGATGTAATTGGACCGGAGGTTTCCGATGCTTTAGAA 120
Db 5869 CTGCAATGTGTGATAGGCGATGTAATTGGACCGGAGGTTTCCGATGCTTTAGAA 5928
Qy 121 CGAAAGCTTCTAGGGGAAGAGATTACTCAAGCTCCACGAGTGCAGCGCTGGGTTA 180
Db 5929 CGAAAGCTTCTAGGGGAAGAGATTACTCAAAATCTCCACGAGTGCAGCGCTGGGTTA 5988
Qy 181 GCTTGCAAAAGGTGCGCATGATGTCCTTCCGAGATCTCAGCTGTGAAAGTAGCA 240
Db 5989 GCTTGCAAAAGGTGCGCATGATGTCCTTCCGAGATCTCAGCTGTGAAAGTAGCA 6048
Qy 241 GTTCTCTCGGCTTATATCTATATTTGGCAAGAAATTTCTGGGGGATCTTTGAAATGAAA 300
Db 6049 GTTCTCTCGGCTTATATCTATATTTGGCAAGAAATTTCTGGGGGATCTTTGAAATGAAA 6108
Qy 301 CTTTGGGCGGCGCAAGTGAAGTGTCTGTTGCCATGCACTTGGACATTTGCAATG 360
Db 6109 CTTTGGGCGGCGCAAGTGAAGTGTCTGTTGCCATGCACTTGGACATTTTGCATG 6168
Qy 361 GATTTCTCCGAAGCACAACTAATCCAAAGCCGTGTTTGTGACGCGTAAAGATGTGA 420
Db 6169 GATTTCTCCGAAGCACAACTAATCCAAAGCCGTGTTTGTGACGCGTAAAGATGTGA 6228
Qy 421 CCGATTGATCTTATATTTTCTGCGCATTTCAATCTTCAAGCTCCGCGCTTTGCAAC 480
Db 6229 CCGATTGATCTTATATTTTCTGCGCATTTCAATCTTCAAGCTCCGCGCTTTGCAAC 6288
Qy 481 CTGCAATGCGGCTGTAGCAAGATGCAAGTGAATGCGTTTCCGGGTTTACATTAAC 540
Db 6289 CTGCAATGCGGCTGTAGCAAGATGCAAGTGAATGCGTTTCCGGGTTTACATTAAC 6348
Qy 541 CTTGAAAGGCGCGTGCATTTGAATGATGTTAGTGAACCTGAATGCTGAAGGTT 600
Db 6349 CTTGAAAGGCGCGTGCATTTGAATGATGTTAGTGAACCTGAATGCTGAAGGTT 6408
Qy 601 TCGGCAAGGTTCTTTTCCAAATGCACTGCTCTAGACATGCAACCGTTTGTGACCA 660
Db 6409 TCGGCAAGGTTCTTTTCCAAATGCACTGCTCTAGACATGCAACCGTTTGTGACCA 6468
Qy 661 TGTTCGATAGTGAAGATCGGCTCTTCCGGAAGATGTTCTTAAGCGGAAGGGG 720
Db 6469 TGTTCGATAGTGAAGATCGGCTCTTCCGGAAGATGTTCTTAAGCGGAAGGGG 6528
Qy 721 GTCATTTGGCGTGGCATTTCCGGAATCGTGTGGCAAGAACTGCTTATGCTGGGGTA 780
Db 6529 GTCATTTGGCGTGGCATTTCCGGAATCGTGTGGCAAGAACTGCTTATGCTGGGGTA 6588
Qy 781 GACGATGTTACATATATGAAGCAAGTATCGTGTGGCAAGCTTTGTGACATGCT 840
Db 6589 GACGATGTTACATATATGAAGCAAGTATCGTGTGGCAAGCTTTGTGACATGCT 6648
Qy 841 TTCAGGAGGCTCTAGTGTGCGGCAAGGCGGAGTGGATGATTTCTCTGCTGCA 900
Db 6649 TTCAGGAGGCTCTAGTGTGCGGCAAGGCGGAGTGGATGATTTCTCTGCTGCA 6708
Qy 901 TTCTGCTTTTCTTCTCTGAGCGTTACGAGCTGTCTTGCATGAGGCGGTTCCCAAT 960
Db 6709 TTCTGCTTTTCTTCTCTGAGCGTTACGAGCTGTCTTGCATGAGGCGGTTCCCAAT 6768
Qy 961 CCGGCGCAGTGCACCTTACTTGTGTCACAAAGCGTCCATATCATGTGAAAGCGGG 1020
Db 6769 CCGGCGCAGTGCACCTTACTTGTGTCACAAAGCGTCCATATCATGTGAAAGCGGG 6828
Qy 1021 CAGCTGCACGAGGTGTCATGCGCTTCAACAGTGTGGCGCTTTGAGAGAC 1080
Db 6829 CAGCTGCACGAGGTGTCATGCGCTTCAACAGTGTGGCGCTTTGAGAGAC 6888
Qy 1081 GGTTTTCATGACGAGATATGTGTGCTGTGCGCTGTGCTATTAATCAGGCGTTGAA 1140
Db 6889 GGTTTTCATGACGAGATATGTGTGCTGTGCGCTGTGCTATTAATCAGGCGTTGAA 6948

QY	1141	TCAGGACACATTAGGTGGGCTCATGACTCTGGCAAAATTTGGCTGAACCGTTTGGGAGG	1200
Db	6949	TCAGGACACATTAGGTGGGCTCATGACTCTGGCAAAATTTGGCTGAACCGTTTGGGAGG	7008
QY	1201	GAGTCCTTCTCTTAGGGATGAGAGGATCTTTCTGGGGACACATCCCTCGGTGGT-AA	1259
Db	7009	GAGTCCTTCTCTTAGGGATGAGAGGATCTTTCTGGGGACACATCCCTCGGTGGTGBA	7068
QY	1260	ACATGGAGTTTTCCTCATGATTGGGACCTATTCAAGCTAATGGGAATAGATCTGGCGGG	1319
Db	7069	ACATGGAGTTTTCCTCATGATTGGGACCTATTCAAGCTAATGGGAATAGATCTGGC-GG	7127
QY	1320	GTTTGGTCCAGTTTTTTGAAAGCGGGTTTATTGAATCTCCGCTTGGTCATCAACGATA	1379
Db	7128	GTTTGGTCCAGTTTTTTGAAAGCGGGTTTATTGAATCTCCGCTTGGTCATCAACGATA	7187
QY	1380	TGAAGAAATCAGGGGATGTGCCCTGAAGGAATCTCAGAACTTCCAGCTCGGATCGCATC	1439
Db	7188	TGAAGAAATCAGGGGATGTGCCCTGAAGGAATCTCAGAACTTCCAGCTCGGATCGCATC	7247
QY	1440	TGAAGTGGTTAAACGGTGTGTGTGAGCCAGACGATATGCCATTTCAAGTCAGGCGCAT	1499
Db	7248	TGAAGTGGTTAAACGGTGTGTGTGAGCCAGACGATATGCCATTTCAAGTCAGGCGCAT	7307
QY	1500	TCAGAGGAAAAAGACAAAAATTAAGATTAAGGCTTTAAGAGCGGATATCTGAACTTTATGA	1559
Db	7308	TCAGAGGAAAAAGACAAAAATTAAGATTAAGGCTTTAAGAGCGGATATCTGAACTTTATGA	7367
QY	1560	TAAAGTGTGTGCATCTGGAATCCGCAAAATCTCAACTCAGGCAATTCCTCGAATCGGA	1619
Db	7368	TAAAGTGTGTGCATCTGGAATCCGCAAAATCTCAACTCAGGCAATTCCTCGAATCGGA	7427
QY	1620	TACCAATATTTTAAAGGACACAGTGAACCAAGCGGTTATTAACAGCCATATGACAGATC	1679
Db	7428	TACCAATATTTTAAAGGACACAGTGAACCAAGCGGTTATTAACAGCCATATGACAGATC	7487
QY	1680	GTCAAAACTCTTCCTGATGACTGAACGAAAAATCTGTGTAGACCATATCTCCGCTTG	1739
Db	7488	GTCAAAACTCTTCCTGATGACTGAACGAAAAATCTGTGTAGACCATATCTCCGCTTG	7547
QY	1740	TGTCTCTATGACGGGATCCGCAAAAGCAGTGTATTGCTTGGACATATGAGTGGCAGATCC	1799
Db	7548	TGTCTCTATGACGGGATCCGCAAAAGCAGTGTATTGCTTGGACATATGAGTGGCAGATCC	7607
QY	1800	GAATGGTAAAGSTCTAGTGCATCAGTTATATACATGGAGGACGACCTCCCAAGACTTT	1859
Db	7608	GAATGGTAAAGSTCTAGTGCATCAGTTATATACATGGAGGACGACCTCCCAAGACTTT	7667
QY	1860	GCGCGTCCCGACAAAAAAGACGATTAATGTCTGCTGGGAGCGCAATTTGCAATCTTT	1919
Db	7668	GCGCGTCCCGACAAAAAAGACGATTAATGTCTGCTGGGAGCGCAATTTGCAATCTTT	7727
QY	1920	CCCCGCGTTTGGCCAGCACCTATTCTGTGCTGCGCTGATTAACGACCAAAATGTTATTCA	1979
Db	7728	CCCCGCGTTTGGCCAGCACCTATTCTGTGCTGCGCTGATTAACGACCAAAATGTTATTCA	7787
QY	1980	ACATGATTGGCTTACAGACGAGAAATGCGGGGGGACCTTCAAACCTCAACCGGCGTGTGA	2039
Db	7788	ACATGATTGGCTTACAGACGAGAAATGCGGGGGGACCTTCAAACCTCAACCGGCGTGTGA	7847
QY	2040	GGATTTTATTTCTGAAGAACTTTTCTTCAAGCACATGACACGCGTAAATGATACCGGAGT	2099
Db	7848	GGATTTTATTTCTGAAGAACTTTTCTTCAAGCACATGACACGCGTAAATGATACCGGAGT	7907
QY	2100	TTACTTGGCGGGTTGCAGTTGTTCTTTCACAGGTGATGGGTGAGGGTGCTATTCAAC	2155
Db	7908	TTACTTGGCGGGTTGCAGTTGTTCTTTCACAGGTGATGGGTGAGGGTGCTATTCAAC	7967
QY	2160	CGCGGTAAACCGCGTGTGCAATTAATCCAAATTTGTGAGAGCATTTTGGCAAGGGGA	2219
Db	7968	CGCGGTAAACCGCGTGTGCAATTAATCCAAATTTGTGAGAGCATTTTGGCAAGGGGA	8027

Oy	2220	TCTCTCGAAGCACTCTTGGAAAGATATACTACCCGAGTAGAAATTAG	2268
Dn	8028	TCTCTCGAAGCACTTTGGAAAGATATACTACCCGACTAGAAATTAG	8076
RESULT 5			
LOCUS	AF242881	194140 bp	DNA circular BCT 26-JUN-2000
DEFINITION	Agrobacterium tumefaciens octopine-type Ti plasmid, complete sequence.		
ACCESSION	AF242881 AF034854 AF034855 AF035413 AF035773 AF039887 AF039888 AF112804 AH006976 M14480 M14762 M80605 M80607 U19620 U43674 U43675 U48718 U60011 X04784		
VERSION	AF242881.1	GI:8572673	
SOURCE			
ORGANISM	Agrobacterium tumefaciens (Rhizobium radiobacter)		
REFERENCE	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.		
AUTHORS	1 (bases 193234 to 194140; 1 to 23692) Barker,R.F., Idler,K.B., Thompson,D.V. and Kemp,J.D.		
TITLE	Nucleotide sequence of the T-DNA region from the Agrobacterium tumefaciens octopine Ti plasmid pTi15955		
JOURNAL	Plant Mol. Biol. 2, 335-350 (1993)		
REMARK	original sequence presented in Genbank Accession Number X00493		
AUTHORS	2 (bases 166443 to 168193) Yanofsky,M.F. and Nester,E.W.		
REFERENCE	Molecular characterization of a host-range-determining locus from Agrobacterium tumefaciens		
TITLE	J. Bacteriol. 168 (1), 244-250 (1986)		
JOURNAL	87008387		
MEDLINE	3759904		
PUBMED	3 (bases 168089 to 173951) Yanofsky,M.F., Porter,S.G., Young,C., Albritton,L.M., Gordon,M.P. and Nester,E.W.		
REFERENCE	The virD operon of Agrobacterium tumefaciens encodes a site-specific endonuclease		
TITLE	Cell 47 (3), 471-477 (1986)		
JOURNAL	87028239		
MEDLINE	3021341		
PUBMED	4 (bases 175278 to 177719) Winans,S.C., Allena,P., Stachel,S.E., McBride,K.E. and Nester,E.W.		
REFERENCE	Characterization of the vira operon of the Agrobacterium Ti plasmid pTiA6		
AUTHORS	Nucleic Acids Res. 15 (2), 825-837 (1987)		
TITLE	87146396		
JOURNAL	MEJNE		
MEDLINE	3547330		
PUBMED	5 (bases 50125 to 54471) Habeeb,L.F., Wang,L. and Winans,S.C.		
REFERENCE	Transcription of the octopine catabolism operon of the Agrobacterium tumor-inducing plasmid pTiA6 is activated by a LysR-type regulatory protein		
AUTHORS	Lyer. Plant Microbe Interact. 4 (4), 379-385 (1991)		
TITLE	9210631		
JOURNAL	MEJNE		
MEDLINE	1799699		
PUBMED	6 (bases 50125 to 54471) Valdivia,R.H., Wang,L. and Winans,S.C.		
REFERENCE	Characterization of a putative periplasmic transport system for octopine accumulation encoded by Agrobacterium tumefaciens Ti plasmid pTiA6		
AUTHORS	J. Bacteriol. 173 (20), 6398-6405 (1991)		
TITLE	9201141		
JOURNAL	MEJNE		
MEDLINE	1655707		
PUBMED	7 (bases 68024 to 76460) Kim,K.S. and Farrand,S.K.		
REFERENCE	Ti plasmid-encoded genes responsible for catabolism of the crown gall opine mannopine by Agrobacterium tumefaciens are homologs of the T-region genes responsible for synthesis of this opine by the plant tumor		
AUTHORS	J. Bacteriol. 178 (11), 3275-3284 (1996)		
TITLE	96236046		
JOURNAL	MEJNE		
MEDLINE	8655509		
PUBMED	8 (bases 39167 to 45466)		
REFERENCE			

AUTHORS
TITLE
Fagua, C. and Winans, S.C.
Localization of OccR-activated and Trar-activated promoters that
express two ABC-type permeases and the trar gene of Ti plasmid
pTiR10

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
9 (bases 22855 to 39243; 96353 to 111409)
Alt-Morbe, J., Stryker, J.L., Fagua, C., Li, P.L., Farrand, S.K. and
Winans, S.C.
The conjugal transfer system of Agrobacterium tumefaciens
octopine-type Ti plasmids is closely related to the transfer system
of an IncP plasmid and distantly related to Ti plasmid vir genes

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
J. Bacteriol. 178 (14), 4248-4257 (1996)
96312368
8763954
10 (bases 54312 to 62806)
Oger, P., Kim, K.-S., Sackett, R.L., Piper, K.R. and Farrand, S.K.
Octopine-type Ti plasmids code for a mannopine-inducible
dominant-negative allele of trar, the quorum-sensing activator that
regulates Ti plasmid conjugal transfer

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
Mol. Microbiol. 27 (2), 277-288 (1998)
98143415
9484884
11 (bases 133963 to 138972)
Kalogeraki, V.S. and Winans, S.C.
Wound-released chemical signals may elicit multiple responses from
an Agrobacterium tumefaciens strain containing an octopine-type Ti
plasmid

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
J. Bacteriol. 180 (21), 5660-5667 (1998)
99009000
9791116
12 (bases 76455 to 98723)
Ly, S.M., Jafri, S. and Winans, S.C.
Mannopine acid and agropine acid catabolism region of the
octopine-type Ti plasmid pTi15955

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
Mol. Microbiol. 31 (1), 339-347 (1999)
99141607
9987134
13 (bases 1 to 194140)
Winans, S.C., Zhu, J., Oger, P.M., Schrammeyer, B., Hooykaas, P.J. and
Farrand, S.K.
Octopine-type Ti plasmid sequence

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
Unpublished
14 (bases 68024 to 76460)
Kim, K.-S.
Direct Submission
Submitted (10-JAN-1995) Kun-Soo Kim, Microbiology, University of
Illinois at Urbana-Champaign, 1201 W. Gregory Dr., Urbana, IL
61801, USA

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
15 (bases 96353 to 111409)
Winans, S.C.
Direct Submission
Submitted (20-DEC-1995) Stephen C. Winans, Microbiology, Cornell
University, Wing Hall, Ithaca, NY 14853, USA

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
16 (bases 39167 to 45466)
Fagua, C.
Direct Submission
Submitted (07-FEB-1996) Clay Fagua, Biology, Trinity University,
715 Stadium Dr., San Antonio, TX 78212, USA

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
17 (bases 54312 to 62806)
Kim, K.-S., Sackett, R.L. and Farrand, S.K.
Direct Submission
Submitted (04-JUN-1996) Crop Sciences, University of Illinois at
Urbana-Champaign, 240 ERM, 1201 W. Gregory Dr., Urbana, IL 61801,
USA

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
18 (bases 186239 to 193239)
Zhu, J. and Winans, S.C.
Direct Submission
Submitted (17-NOV-1997) Section of Microbiology, Cornell
University, Wing Hall, Ithaca, NY 14853, USA

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
19 (bases 76455 to 98723)
Winans, S.C. and Ly, S.

TITLE
JOURNAL
Direct Submission
Submitted (20-NOV-1997) Microbiology, Cornell University, Wing
Hall, Ithaca, NY 14853, USA

REFERENCE
AUTHORS
TITLE
20 (bases 170209 to 175283)
Zhu, J. and Winans, S.C.
Direct Submission
Submitted (25-NOV-1997) Section of Microbiology, Cornell
University, Wing Hall, Ithaca, NY 14853, USA

REFERENCE
AUTHORS
TITLE
21 (bases 133963 to 138972; 176911 to 186244)
Stryker, J.L., Mantis, N.J., Kalogeraki, V.S. and Winans, S.C.
Direct Submission
Submitted (23-DEC-1997) Microbiology, Cornell University, Ithaca,
NY 14853, USA

REFERENCE
AUTHORS
TITLE
22 (bases 22855 to 39243)
Winans, S.C.
Direct Submission
Submitted (03-AUG-1998) Microbiology, Cornell University, Wing
Hall, Ithaca, NY 14853, USA

REFERENCE
AUTHORS
TITLE
23 (bases 62801 to 66985)
Oger, P.M.
Direct Submission
Submitted (29-DEC-1998) Crop Sciences, University of Illinois at
Urbana-Champaign, 361 ERM, 1201 W. Gregory Dr., Urbana, IL 61801,
USA

REFERENCE
AUTHORS
TITLE
24 (bases 110822 to 133968)
Zhu, J. and Winans, S.C.
Direct Submission
Submitted (18-AUG-1999) Section of Microbiology, Cornell
University, Wing Hall, Ithaca, NY 14853, USA

REFERENCE
AUTHORS
TITLE
25 (bases 1 to 194140)
Zhu, J., Oger, P.M., Schrammeyer, B., Hooykaas, P.J., Farrand, S.K. and
Winans, S.C.
Direct Submission
Submitted (07-MAR-2000) Microbiology, Cornell University, Wing
Hall, Ithaca, NY 14853, USA

COMMENT
On or before Jun 20, 2000 this sequence version replaced
gi:3377769, gi:1215729, gi:1381799, gi:797330, gi:2665704,
gi:2749889, gi:15738260, gi:2773254, gi:2773255, gi:2773256,
gi:154820, gi:154827, gi:2662524, gi:39122, gi:2687604, gi:154767,
gi:154773, gi:6689867.

FEATURES
source
1. 194140
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:358"
/plasmid="Ti"
/note="extrachromosomal
octopine-type"
13244..13618
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:358"
/insertion_seg="ybd4"
13636..14181
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:358"
/insertion_seg="ybe"
/note="similar to IS66"
14183..14666
/organism="Agrobacterium tumefaciens"

Query Match 98.4%; Score 2230.6; DB 1; Length 194140;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2258; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 1 ATGTACGCTTACCTCTCTGATTAACCGATCGATCACTCCCAACAAATGCGGAT 60
|||||
DB 4902 ATGTACGCTTACCTCTCTTATTAACCGATCGATCACTCCCAACAAATGCGGAT 4961
|||||
QY 61 CTGACAAATGCGATTAAGCGGATGAATTGACCGGAGGTTTCGATCCTTTAGAA 120
|||||
DB 4962 CTGACAAATGCGATTAAGCGGATGAATTGACCGGAGGTTTCGATCCTTTAGAA 5021
|||||

QY 121 CGAAGCTTCTAGGGGAAAGAGATTACTCAAAAGCTCACAGAGTGCAGCGTGGGTTA 180
DB 5022 CGAAGAGCTTCTAGGGGAAAGAGATTACTCAAAATCTCCACCGAGTGCAGCGTGGGTTA 5081
QY 181 GCTTGGCAAAAGGCTGGCCGATGGTGGCTTCCCGAGATCTCAGCTGGTGAAGTAGCA 240
DB 5082 GCTTGGCAAAAGGCTGGCCGATGGTGGCTTCCCGAGATCTCAGCTGGTGAAGTAGCA 5141
QY 241 GTTCTCTCCGCTTATATCTATATATGGCAAAAGAAATCTGGGGGGGATACTTGAATGGAA 300
DB 5142 GTTCTCTCCGCTTATATCTATATATGGCAAAAGAAATCTGGGGGGGATACTTGAATGGAA 5201
QY 301 CCTTGGGCGCGGCAACAGTAGTGTCTCGTGGCCATCGACTTGAACACATTTTGCATG 360
DB 5202 CCTTGGGCGCGGCAACAGTAGTGTCTCGTGGCCATCGACTTGAACACATTTTGCATG 5261
QY 361 GATTTCTCCGAGACACAATCTAATCCAAAGCCCTGTTTTGCTGAGCGGTAAAGATGTGA 420
DB 5262 GATTTCTCCGAGACACAATCTAATCCAAAGCCCTGTTTTGCTGAGCGGTAAAGATGTGA 5321
QY 421 CCGATTGATTTAGTCAATTTCTGGGCAATTTCAATCTCTAAGACTGCCGCTTTCCAAAC 480
DB 5322 CCGATTGATTTAGTCAATTTCTGGGCAATTTCAATCTCTAAGACTGCCGCTTTCCAAAC 5381
QY 481 CTGCGCAATGCGGCTGATGAGAAATGGCAAGATGAATGCGTTACCGGGTTTACCATPACC 540
DB 5382 CTGCGCAATGCGGCTGATGAGAAATGGCAAGATGAATGCGTTACCGGGTTTACCATPACC 5441
QY 541 CTTGAAAGGGGCGGTCCATTTGACATGTGATGCTTATGTGGAACCTGATGTGGAAGGT 600
DB 5442 CTTGAAAGGGGCGGTCCATTTGACATGTGATGCTTATGTGGAACCTGATGTGGAAGGT 5501
QY 601 TCGGCAAGTTCCTTTCCAAATCGAATCTGCTCTAAGACTGCAAGCCGTTTTTGAACCA 660
DB 5502 TCGGCAAGTTCCTTTCCAAATCGAATCTGCTCTAAGACTGCAAGCCGTTTTTGAACCA 5561
QY 661 TGTTCGATAGTGAAGGATCGGCTTCTTCCGAGAGATGTTCCTAAGCCGAAAGTGGCG 720
DB 5562 TGTTCGATAGTGAAGGATCGGCTTCTTCCGAGAGATGTTCCTAAGCCGAAAGTGGCG 5621
QY 721 GTCATTTGCGCTGGCATTTCCGAGCTGTGTGTGCAAAAGCAATGCTTCAATGCTGGGTTA 780
DB 5622 GTCATTTGCGCTGGCATTTCCGAGCTGTGTGTGCAAAAGCAATGCTTCAATGCTGGGTTA 5681
QY 781 GAGGATGTTCAATATATGAAGCAAGTATGCTGTGTGAGGCAAGCTTTGTGCAATGCT 840
DB 5682 GAGGATGTTCAATATATGAAGCAAGTATGCTGTGTGAGGCAAGCTTTGTGCAATGCT 5741
QY 841 TTCAAGGAGCGCTCCTAGTGTGTGTGAGGCAAGTGGGCGGATGCGATTTCTCTGCTGCA 900
DB 5742 TTCAAGGAGCGCTCCTAGTGTGTGTGAGGCAAGTGGGCGGATGCGATTTCTCTGCTGCA 5801
QY 901 TTCTGCTGTTTTTCTTCTCTGAGCGTTACGCGCTGTCTTGATGAGCGGTTCCCAAT 960
DB 5802 TTCTGCTGTTTTTCTTCTCTGAGCGTTACGCGCTGTCTTGATGAGCGGTTCCCAAT 5861
QY 961 CCGGCGCACTGCAACTTACTTGTGTGTTACCAAGCGTCCAAATCAATGTGAAAGCCGGG 1020
DB 5862 CCGGCGCACTGCAACTTACTTGTGTGTTACCAAGCGTCCAAATCAATGTGAAAGCCGGG 5921
QY 1021 CAGTGCACGGAAGCTGTTCAATCGGTTTACCAAGGTTGCGGTGCTTGAAGGAC 1080
DB 5922 CAGTGCACGGAAGCTGTTCAATCGGTTTACCAAGGTTGCGGTGCTTGAAGGAC 5981
QY 1081 GGTTCATGAGCGAGATATGTGTGTGGCTTGCGCTGTATTAATCAAGCCCTTGA 1140
DB 5982 GGTTCATGAGCGAGATATGTGTGTGGCTTGCGCTGTATTAATCAAGCCCTTGA 6041
QY 1141 TCAGGACACATTAAGTGGGCTCATGACTCTGCGCAATTTGGCTGAACGTTTTGGGAGG 1200
DB 6042 TCAGGACACATTAAGTGGGCTCATGACTCTGCGCAATTTGGCTGAACGTTTTGGGAGG 6101

QY 1201 GAGTCTTCTCTCAGGATAGAGAGATCTTTCTGGGCAACATCTCTCGTGTGT -AA 1259
DB 6102 GAGTCTTCTCTCAGGATAGAGAGATCTTTCTGGGCAACATCTCTCGTGTGTGTGA 6161
QY 1260 ACATGAGATTTTCTCATATATGGGACCTTATCAAGCTTAATGGGAAATAGATCTGGCGGG 1319
DB 6162 ACATGAGATTTTCTCATATATGGGACCTTATCAAGCTTAATGGGAAATAGATCTGGC-GG 6220
QY 1320 GTTTGGTCCAGTTTTTGAAGGCGGTTTTTATGAGATCTCTCGCTTGTGATCAACGGATA 1379
DB 6221 GTTTGGTCCAGTTTTTGAAGGCGGTTTTTATGAGATCTCTCGCTTGTGATCAACGGATA 6280
QY 1380 TGAAGAAATCAAGCGGATGTGCTTGAAGAAATCTCAGAACTTCCAGTGTGCATC 1439
DB 6281 TGAAGAAATCAAGCGGATGTGCTTGAAGAAATCTCAGAACTTCCAGTGTGCATC 6340
QY 1440 TGAAGTGTAAAGGTTGTCTGTGAGCCAGGCGCATATGCCATGTTCAAGTCAAGGCGAT 1499
DB 6341 TGAAGTGTAAAGGTTGTCTGTGAGCCAGGCGCATATGCCATGTTCAAGTCAAGGCGAT 6400
QY 1500 TCAGAGGAAAGACAAATAATTAAGCTTAAGACGCGGATATCTGAACCTTATGA 1559
DB 6401 TCAGAGGAAAGACAAATAATTAAGCTTAAGACGCGGATATCTGAACCTTATGA 6460
QY 1560 TAAAGTGTGTCAATCTGCACTGCAAAATATCAACTCAGGCAATGCTGATGCA 1619
DB 6461 TAAAGTGTGTCAATCTGCACTGCAAAATATCAACTCAGGCAATGCTGATGCA 6520
QY 1620 TACCAATATTTTCAAGGCAACAGTGAACCAAGCGGTGTTAACAGCAATATGA 1679
DB 6521 TACCAATATTTTCAAGGCAACAGTGAACCAAGCGGTGTTAACAGCAATATGA 6580
QY 1680 GTCAAAACTCTTCTGATGATGTAAGAAATTTCTGTTAGACATATCTCCGCTTGG 1739
DB 6581 GTCAAAACTCTTCTGATGATGTAAGAAATTTCTGTTAGACATATCTCCGCTTGG 6640
QY 1740 TGTCTCATGAGAGGATGCAAAAGCAAGTATGCTGAGACTATGAGTGCAGAGATCC 1799
DB 6641 TGTCTCATGAGAGGATGCAAAAGCAAGTATGCTGAGACTATGAGTGCAGAGATCC 6700
QY 1800 GAATGTTAAAGTCTAGTGTCTATCAATTAATATGGAGAGAGACTCCCAAGCTGTT 1859
DB 6701 GAATGTTAAAGTCTAGTGTCTATCAATTAATATGGAGAGAGACTCCCAAGCTGTT 6760
QY 1860 GGGGTCGCCGACAAATAAGAGGATTAATGTCTGTGCGGAGCGCAATTTGAGATCTTT 1919
DB 6761 GGGGTCGCCGACAAATAAGAGGATTAATGTCTGTGCGGAGCGCAATTTGAGATCTTT 6820
QY 1920 CCGGCGGTTTTGCCAGACCTATTTCTGCTGCGCTGATTAAGACCAAAATGTTATCA 1979
DB 6821 CCGGCGGTTTTGCCAGACCTATTTCTGCTGCGCTGATTAAGACCAAAATGTTATCA 6880
QY 1980 ACATGATTTGGCTTACAGAGAAATGCGGCGGAGCTTTCAACTCAACCGGCGGTGA 2039
DB 6881 ACATGATTTGGCTTACAGAGAAATGCGGCGGAGCTTTCAACTCAACCGGCGGTGA 6940
QY 2040 GGAATTTTATTTGAGAACTTTCTTTCAAGCACTGACACCGGCTAATGATACCGGAGT 2099
DB 6941 GGAATTTTATTTGAGAACTTTCTTTCAAGCACTGACACCGGCTAATGATACCGGAGT 7000
QY 2100 TTAATTTGGCGGTTGAGTTGTTCTTCAAGGTGATGAGTGAAGGTTGATTTCAAG 2159
DB 7001 TTAATTTGGCGGTTGAGTTGTTCTTCAAGGTGATGAGTGAAGGTTGATTTCAAG 7060
QY 2160 CCGGTGTAACCGGCTGTGTGCAATTAATCCCAATTTGTGAGGCAATTTTGGCAAGGCAA 2219
DB 7061 CCGGTGTAACCGGCTGTGTGCAATTAATCCCAATTTGTGAGGCAATTTTGGCAAGGCAA 7120
QY 2220 TCCCTCGAACACTCTTGAAGAGATTAATCAACCGGATGAGAAATTTAG 2268
DB 7121 TCCCTCGAACACTCTTGAAGAGATTAATCAACCGGATGAGAAATTTAG 7169

RESULT 6
TIMPS2
LOCUS
DEFINITION
T1 plasmid (A. tumefaciens octopine strain) tms1 gene, complete
ACCESSION
K02554
VERSION
K02554.1 GI:154747
KEYWORDS
2 of 2
SEGMENT
plasmid T1
SOURCE
plasmid T1
ORGANISM
1 (bases 1 to 2591)
REFERENCE
Klee, H., Montoya, A., Horodyski, F., Lichtenstein, C., Garfinkel, D.,
Fuller, S., Flores, C., Peschon, J., Nester, E. and Gordon, M.
Nucleotide sequence of the tms genes of the PTiA6NC octopine Ti
plasmid: two gene products involved in plant tumorigenesis
Proc. Natl. Acad. Sci. U.S.A. 81 (6), 1728-1732 (1984)
JOURNAL
MEDLINE
PUBMED
84170374
COMMENT
6584906
Original source text: Plasmid T1 (clone: PTiA6NC).
The tms1 protein has a region which showed significant homology to
the adenine binding region of p-hydroxybenzoate hydroxylase from
P. fluorescens. It seems that the tms1 protein binds adenine
either as substrate or cofactor.
The initiation and termination sites of transcription have not yet
been identified [1].
FEATURES
source
Location/Qualifiers
1..2591
/organism="Plasmid T1"
/mol_type="genomic DNA"
/specific_host="octopine"
/db_xref="taxon:2512"
/clone="PTiA6NC."
/plasmid="Plasmid T1"
160..2427
/gene="tms1"
160..2427
/gene="tms1"
/codon_start=1
/transl_table=1
/protein_id="AA92550.1"
/db_xref="GI:154750"
/translation="MSASPLDNOCDLPTKWDLTMDVKADLDRVSDAFLERAS
RGRITQISTECAGLACKRLADGRPEISAGSKVAVLSYITIGELIKRILIESRPW
ARATVGLVAIDLAPCMDFESQQLQALFLSGKRCAPLIDSHFVAISIKTAGPRT
LMPVLENGTMKCVGTFTTLEGAVPDMVAGRNLMKSGASPTIDLDLYRPF
DQSDSGRIGFPEDVAPKRPVAVIGAGISGLVAVNELLAGVDPVTIYASDVGKL
WSHAFDADSVVAEMGMRPPAPFLPFELRYGSLSMRPPENPCTVDVTIYVGVQV
YMKAGQDLPKLEHRYNMGRAFLLKDFPERDVLASPAVITDALKSGDIRAHDSMQ
IMNFRGRESFSGIERIFLGHFPGESTWSPFHDWLEFLMGISGGFGVPESGFI
ELRLVINGEENORMCPBGISELPRRIASEVNGVSVQRICHVQVRAIQEKTKIK
IRKSGISLIDVVTVSGLANIQLRHCLTCDNIFQAPVNOAVDSHMTGSKLFLM
TERKFWLDHILPESCVLMGIAKAVYCLDYEPDOPNGKGLVLIYTWEDSHKLAVPD
KKERLLDLDAISRFPFAFOHLFPACADYDQNVIOHDMLTDENAGAKRIARBGDF
YSEELFFQALDPTANDTNGVLIAGSCSFTGGMWGAIQYACNAVCAIINHCGGILANGN
PLEHSWKRYNRYTRN"
BASE COUNT
ORIGIN
654 a 570 c 656 g 711 t
About 250 bp after segment 1.
Query Match 98.1%; Score 2225.8; DB 1; Length 2591;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2255; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

DB 280 CGAAGAGCTTCTAGGGGAAGAGAGATTACTCAAAATCTCCACCGAGTGCAGCGCTGGGTTA 339
QY 181 GCTTGCAGAAAGCTGGCCGATGTGCTTCCCGAATCTCAGCTGTGGAAAGTAGCA 240
DB 340 GCTTGCAGAAAGCTGGCCGATGTGCTTCCCGAATCTCAGCTGTGGAAAGTAGCA 399
QY 241 GCTTCTCCGCTATATCTATATATGCAAGAAATCTGAGGGGGGATGACTGTAATGAAA 300
DB 400 GTTCTCTCCGCTATATCTATATATGCAAGAAATCTGAGGGGGGATGACTGTAATGAAA 459
QY 301 CTTGGGCGCGGCAACAGTGAAGTGTCTGTTCCATCGACTTGACACCAATTTGCATG 360
DB 460 CTTGGGCGCGGCAACAGTGAAGTGTCTGTTCCATCGACTTGACACCAATTTGCATG 519
QY 361 GATTTCTCCGAGACCACTAATCCAGCCCTGTTTGTGTGCGGTAAGATGTGA 420
DB 520 GATTTCTCCGAGACCACTAATCCAGCCCTGTTTGTGTGCGGTAAGATGTGA 579
QY 421 CGATTTGATCTTATGATCTTCTGAGCAATTTCAATCTTAAAGTCCGCTTTGCAAC 480
DB 580 CGATTTGATCTTATGATCTTCTGAGCAATTTCAATCTTAAAGTCCGCTTTGCAAC 639
QY 481 CTGCAATGCGCGTGTACAGAGATGACAGATGAATGCGTTAACCGGTTTACCATACC 540
DB 640 CTGCAATGCGCGTGTACAGAGATGACAGATGAATGCGTTAACCGGTTTACCATACC 699
QY 541 CTTGAAGGAGCGGTGCAATTTGACATGTGATGCTTAAAGCTGTAAGAGGT 600
DB 700 CTTGAAGGAGCGGTGCAATTTGACATGTGATGCTTAAAGCTGTAAGAGGT 759
QY 601 TGGGCAAGTTCCTTCCCAATCCGCTTCTAGACATGAGACCGTCTTTTGACCA 660
DB 760 TGGGCAAGTTCCTTCCCAATCCGCTTCTAGACATGAGACCGTCTTTTGACCA 819
QY 661 TGTTCGATATGAGACGATTCGCTTCTTCCGAGAGATGTTCTTAAGCGAAAGTGGCG 720
DB 820 TGTTCGATATGAGACGATTCGCTTCTTCCGAGAGATGTTCTTAAGCGAAAGTGGCG 879
QY 721 GTGATTTGAGCGTGGCATTTCCGAGCTGTGTGGGCAACGAATGCTTATGCTGGGTA 780
DB 880 GTGATTTGAGCGTGGCATTTCCGAGCTGTGTGGGCAACGAATGCTTATGCTGGGTA 939
QY 781 GACGATTTACAAATATATGAAGCAAGTATCGTGTGAGGCAAGCTTTGTCACATGCT 840
DB 940 GACGATTTACAAATATATGAAGCAAGTATCGTGTGAGGCAAGCTTTGTCACATGCT 999
QY 841 TTCAGGAGCGCTCTAGTGTGCGGCGAAATGGGGGCGATGATTTCTCTCGTGTGA 900
DB 1000 TTCAGGAGCGCTCTAGTGTGCGGCGAAATGGGGGCGATGATTTCTCTCGTGTGA 1059
QY 901 TTCTGCTTTTCTTCTCTCGAGCGTTACGCGCTGTCTTCGATGAGCGCTTCCCAAT 960
DB 1060 TTCTGCTTTTCTTCTCTCGAGCGTTACGCGCTGTCTTCGATGAGCGCTTCCCAAT 1119
QY 961 CCCGCGACGTGACACTTACTTGTCTTCAAGGCGTCAATCATGTGGAAGCCGGG 1020
DB 1120 CCCGCGACGTGACACTTACTTGTCTTCAAGGCGTCAATCATGTGGAAGCCGGG 1179
QY 1021 CAGCTGCGACGAGGTGTTCCATCGCGTTCACAGGTTGAGGCGGCTTCTGAAGGAC 1080
DB 1180 CAGCTGCGACGAGGTGTTCCATCGCGTTCACAGGTTGAGGCGGCTTCTGAAGGAC 1239
QY 1081 GGTTCATGAGCGAGATATGTGTTGGCTTGGCTGTGCTATTAATCAGGCTTGAAA 1140
DB 1240 GGTTCATGAGCGAGATATGTGTTGGCTTGGCTGTGCTATTAATCAGGCTTGAAA 1299
QY 1141 TGAAGACATTAAGTGGGCTATGACTCTGCGAAATTTGGCTGAACGCTTTCCGAGG 1200
DB 1300 TGAAGACATTAAGTGGGCTATGACTCTGCGAAATTTGGCTGAACGCTTTCCGAGG 1359
QY 1201 GAGTCCTTCTTCAAGGAGTAAAGAGATCTTTCGGGACACATCTCTGCTGGTGT -AA 1259

Db 1360 GAGTCCCTTCTTCTCAGGATAGAGAGATCTTTCTGGGACACATCTCCTGGNGGAA 1419
 Qy 1260 ACATGAGTTTTTCTCTCATGATTGGGACCTATTCAAGTATATGGGAATAGATCTGGCGG 1319
 Db 1420 ACATGAGTTTTTCTCTCATGATTGGGACCTATTCAAGTATATGGGAATAGATCTGGCG- 1478
 Qy 1320 GTTTGGTCCAGTTTTTGAAGCGGGTTTATAGATCCTCCGCTGGTCACTCAACGGATA 1379
 Db 1479 GTTTGGTCCAGTTTTTGAAGCGGGTTTATAGATCCTCCGCTGGTCACTCAACGGATA 1538
 Qy 1380 TGAAGAAATACACCGGATGTCCTTGAAGGAATCTCGAACTTCCAGTCCGATCCGATC 1439
 Db 1539 TGAAGAAATACACCGGATGTCCTTGAAGGAATCTCGAACTTCCAGTCCGATCCGATC 1598
 Qy 1440 TGAAGTGTGTTAAGGTTGTCTGTGAGCCGCGATTTGCGATTTCAATCGAGGGGAT 1499
 Db 1599 TGAAGTGTGTTAAGGTTGTCTGTGAGCCGCGATTTGCGATTTCAATCGAGGGGAT 1658
 Qy 1500 TCAGAGGAAAGACAAAATTAAGATTAAGGCTTAAGAGCGGGATATCTGAACCTTTATGA 1559
 Db 1659 TCAGAGGAAAGACAAAATTAAGATTAAGGCTTAAGAGCGGGATATCTGAACCTTTATGA 1718
 Qy 1560 TAAAGTGTGTGTCATCTGAGCTCCGAAATATCAACTGAGGATTCCTGACATGCGA 1619
 Db 1719 TAAAGTGTGTGTCATCTGAGCTCCGAAATATCAACTGAGGATTCCTGACATGCGA 1778
 Qy 1620 TACCAATATTTTTCAGGCACCACTGAAACCAAGCGGTTGAATACAGCCATATGACAGATC 1679
 Db 1779 TACCAATATTTTTCAGGCACCACTGAAACCAAGCGGTTGAATACAGCCATATGACAGATC 1838
 Qy 1680 GTCAAACTCTTCCTGATGATGATGAAAGAAATCTGGTTAGACCATATCTCCCGCTTG 1739
 Db 1839 GTCAAACTCTTCCTGATGATGATGAAAGAAATCTGGTTAGACCATATCTCCCGCTTG 1898
 Qy 1740 TGTCTCATGAGCGGATTCGCAAAAGCAGTGTATTCCTGCTGATGATGATGCGAGATCC 1799
 Db 1899 TGTCTCATGAGCGGATTCGCAAAAGCAGTGTATTCCTGCTGATGATGATGCGAGATCC 1858
 Qy 1800 GAATGCTAAAGGTTCTAGTCTCATGATTAATCATGGGAGAGACCTCCCAACAGCTGT 1859
 Db 1959 GAATGCTAAAGGTTCTAGTCTCATGATTAATCATGGGAGAGACCTCCCAACAGCTGT 2018
 Qy 1860 GGGCGTCCCGACAAAAGAGGATATGCTGCTCGGAGCGCAATTCGAGATCTTT 1919
 Db 2019 GGGCGTCCCGACAAAAGAGGATATGCTGCTCGGAGCGCAATTCGAGATCTTT 2078
 Qy 1920 CCCGGCGTTTCCCGACCACTATTTCTGCTGCGCTGATTAACGACCAAAATGTTATCA 1979
 Db 2079 CCCGGCGTTTCCCGACCACTATTTCTGCTGCGCTGATTAACGACCAAAATGTTATCA 2138
 Qy 1980 ACATGATGCTTACACAGAGAAATGCGGGGAGGCTTTCAAACTCAACCGGCGTGGTGA 2039
 Db 2139 ACATGATGCTTACACAGAGAAATGCGGGGAGGCTTTCAAACTCAACCGGCGTGGTGA 2198
 Qy 2040 GATTTTATTCGAAAGACCTTTCTTCAAGCACTGACACCGGCTATGATCCGAGT 2099
 Db 2199 GATTTTATTCGAAAGACCTTTCTTCAAGCACTGACACCGGCTATGATCCGAGT 2258
 Qy 2100 TTAATTGCGCGGTTGCAAGTTTCTTCAAGGATGATGAGTGGAGGCTGCTATTTCAGAC 2159
 Db 2259 TTAATTGCGCGGTTGCAAGTTTCTTCAAGGATGATGAGTGGAGGCTGCTATTTCAGAC 2318
 Qy 2160 CGCGTGAAGCGGCTCTGTGCAATTTATCCAAATTTGAGAGCATTTTGGCAAGGGCAA 2219
 Db 2319 CGCGTGAAGCGGCTCTGTGCAATTTATCCAAATTTGAGAGCATTTTGGCAAGGGCAA 2378
 Qy 2220 TCCCTCGAAGCACTCTTGAAGAGATTAATACTACCGAGTGAAGAAATTAG 2268
 Db 2379 TCCCTCGAAGCACTCTTGAAGAGATTAATACTACCGAGTGAAGAAATTAG 2427

RESULT 7
 AE007927

LOCUS AE007927 9931 bp DNA linear BCT 18-DEC-2001
 DEFINITION Agrobacterium tumefaciens str. C58 plasmid Ti, section 5 of 20 of
 the complete sequence.
 ACCESSION AE007927 AE007871
 VERSION AE007927.2 GI:15163472
 KEYWORDS Agrobacterium tumefaciens str. C58 (Cereon)
 SOURCE Agrobacterium tumefaciens str. C58 (Cereon)
 ORGANISM Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 REFERENCE 1 (bases 1 to 9931)
 AUTHORS Hinkle, G., Slater, S.C. and Goodner, B.
 TITLE Complete Genome Sequence of Agrobacterium tumefaciens C58
 (Rhizobium radiobacter C58), the Causative Agent of Crown Gall
 Disease in Plants
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 9931)
 AUTHORS Hinkle, G., Slater, S.C. and Goodner, B.
 TITLE Direct Submission
 JOURNAL Submitted (14-AUG-2001) Cereon Genomics, 45 Sidney Street,
 Cambridge, MA 02139, USA
 COMMENT On Oct 19, 2001 this sequence version replaced gi:15163470.
 FEATURES
 source
 1. .9931
 /organism="Agrobacterium tumefaciens str. C58 (Cereon)"
 /mol_type="genomic DNA"
 /strain="C58"
 /db_xref="taxon:181661"
 /plasmid="Ti"
 240. .905
 /gene="AGR_pti_bx83"
 240. .905
 /gene="AGR_pti_bx83"
 /note="putative indole-3-lactate synthase"
 /codon_start=1
 /product="AGR_pti_bx83p"
 /protein_id="AK90967.1"
 /db_xref="GI:15163471"
 /translation="MYSRPIPIIINSSNIQDREIKLYVRHTEIAYSPADQDIPA
 QRSWNSIINTDVIDPAIDEVKRCVACLEPGPAGTIPANTILNLSHTYVCSFOAM
 RKVAHRRFYDGVSDGCVISTVPEVAGITKEITMRSMHNNVCONTSENTHDLYAF
 LPTSLQNPFSFHKICDSEFLAPSRVDPCEVETIIAVGKALFHDNGLKPKPRCGQWMLP
 PSC"
 complement (1410. .2813)
 /gene="AGR_pti_48"
 complement (1410. .2813)
 /gene="AGR_pti_48"
 /note="(AJ237588) IAH"
 /codon_start=1
 /transl_table=11
 /product="AGR_pti_48p"
 /protein_id="AK90968.1"
 /db_xref="GI:15163472"
 /translation="MVPITSLAQTLERLRKDYSCLEVETLIARCOAKPLNALIAT
 DMDGLRSARKIDIRHGNAGLGLGIPCEPANIATGIFPSATPLINLPIPSRV
 AERLPSAGALPGASGMMHLSRGITNNVATGAVRPMNPISLPGSSGVAVAASR
 LMIGIGITDGTASVRLPALCGVGRPTLARIPIRRIIPVSTRTAGIINACVADV
 IILDQYISGASAKISMPKLGKRIGLPTTFYDDLDADVAFAETIRLIANGVTFV
 EADIPLHEINDEISASLPALYEPHALKYLDFVGTFSFSDVIKGIISPDVANIISA
 QIDGHOISNDEYLAROSFRPLQATYRVNFRLYQDALIFTPAPLAAKAIGDESSVI
 HNSGMNTEFKIYVRNVDPSNNAGLPGLSIPACLTPLRPVGMETIDGLAGSDHRLAIG
 ALEKXINPSPFDAPN"
 3151. .5418
 /gene="AGR_pti_49"
 3151. .5418
 /gene="AGR_pti_49"
 /note="(AF126446) tryptophan monooxygenase"
 /codon_start=1
 /transl_table=11
 /product="AGR_pti_49p"
 /protein_id="AK90969.1"
 /db_xref="GI:15163473"

/translation="MSASALLDNOCHSTKMDLINVDADEHRSVDSAFSEAS
RGRITQISEGSCAGLACRLADRPFEISAGEKVALSAIYVKGEDILRIESEP
ARARVGLVAILDLAPFCMDPSEAOLOTLFLISGRKASSDLSPVAILSTIKTAR
LOMPYKCTTKTKVTGTLTLEAVPFDMAVGNMLKKSAGSFPTIDLLHYRSPF
DQDSGRIGFPEIDVPRKVALIAGISGLVASELHAGVDVITVYASDRVGL
WSHAFKAPSVVAMGMRPPPAACFELERGLSMPFPMPGTVDNLVYOGRL
YMWKAGQOPKLFHRVYSGWRAFLKDFHGDIVLAPVAILTOLKSGDIRAHDSQ
TMNRFRESFSAIERIFLGTHPGGEWSPHMDLFLMIGSGGFGPFPESGFI
EILRVINGEENORMCSGISLPRIATOVVNGVSQRIHVQRAIEKERTIK
IRKSGISELYDKVYVTSGLANTOLRHCTDITTPAPVNOAVDNSHMTGSLFL
TERFMDLHILPSCVLMGDIKAVYCLDYEPQDNGGLVILSYTWDSDHKLAVD
KSRFCLRLDAISRFPAPQHLVPACADYDQVNOVQDMLTDENAGAFPLNGBEF
YSELFEOALDMTDTGVYLAGSCSFYTGVMVEGAIQTACNAVCAIITHNGILLAKON
PLEHSMKRYVYRNRN"
6089..6838
/gene="AGR_pt1_50"
6089..6838
/gene="AGR_pt1_50"
/note="Probable tmr gene"
/codon_start=1
/translation="AGR_pt1_50p"
/product="AGR_pt1_50p"
/protein_id="AAK90970.1"
/db_xref="GI:15163474"
/translation="MYSRYRKLMIDLIFPTCTGKTSTAVLAOQTGLPVLSDRV
QCCPOLSTGSGRPVEELKGTSLYLDRLVGLIYAKAHRIMEVEVYVYEAHGL
ILGGGSLIKCMAQSVSADPRMILRIELADEBPTMVAARVYQMLRPAAGLSI
IOELVDLKEPRILPIKEIDGTRYAMLFPASQNTSDMLLDLADMEDLHIGIAOE
YLIHARRQEQKFRVNAAYDFEGHFGMY"
7354..7914
/gene="AGR_pt1_51"
7354..7914
/gene="AGR_pt1_51"
/note="Probable 6a gene"
/codon_start=1
/translation="AGR_pt1_51p"
/product="AGR_pt1_51p"
/protein_id="AAK90971.1"
/db_xref="GI:15163475"
/translation="WTRQWLVPECLSHKDDQLODELSELSKAYRRKFOQDLHTKSGD
IIDPSEFLVYLDENYRICRQSMVLSNAPDGLIATLEPVSQSTFQVREOLA
FGSDGRINYSRNEHSSYFLTIQASNEFHVGVANTFQSKDIWRKMPDASQPLD
YHLIIVGCSAFLEBALDVEDSGAV"
complement (8397..9020)
/gene="AGR_pt1_bx9p"
complement (8397..9020)
/gene="AGR_pt1_bx9p"
/note="tm1"
/codon_start=1
/translation="AGR_pt1_bx9p"
/product="AGR_pt1_bx9p"
/protein_id="AAK90972.1"
/db_xref="GI:15163476"
/translation="MTVANQVRLTLITRTGEMOSRLSEQARTFGALLPETVYFOPS
AIRLGFEDDEYILROELVYVLRREDIARCALRHLPSNASSNGIMATLIPWMDA
RLINRYMOERCCDGLNHYQGBHTNOFLAIIINPSCFAPRGDIVINNEYGYANGG
NYTEBESDDDEDEBDEBTGAEIRDSQTNLINTPIALGSHLSA"
BASE COUNT 2639 a 2177 c 2364 g 2751 t
ORIGIN
Query Match 90.1% Score 2044; DB 1; Length 9931;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2141; Conservative 0; Mismatches 125; Indels 2; Gaps 2;

Db 3271 CGAAGAGCTTCTAGGGGAGAGAGATTACTCAATCTCCGGCAGTGCAGCGTGGTTA 3330
Qy 181 GCTTGCAAAAAGGTGGCCGATGTCCTCCGAGATCTACCTGGTGAAGATGCA 240
Db 3331 GCTTGCAAAAAGGTGGCCGATGTCCTCCGAGATCTACCTGGTGAAGATGCA 3390
Qy 241 GTTCTCCGCTATATCTATATGGCAAGAAATCTGGGGGGGATCTTGAATGCAA 300
Db 3391 GCCCTCCGCTTACATCTATGTTGGCAAGAAATCTGGGGGGGATCTTGAATGCAA 3450
Qy 301 CTTGGGCGCGGCAACAGTAGTGTCTGTTGCATCGACTTGAACACATTTGCAT 360
Db 3451 CTTGGGCGCGGCAACAGTAGTGTCTGTTGCATCGACTTGAACACATTTGCAT 3510
Qy 361 GATTTCCGAGACACAACTATCCAGCCCTGTTTTGCTGAGCGGTAAAGATGCA 420
Db 3511 GATTTCCGAGACACAACTATCCAGCCCTGTTTTGCTGAGCGGTAAAGATGCA 3570
Qy 421 CCGATTGATCTTGTGATTTGCGGCATTTCATCTCTAAGATCCGGCTTCCAAAC 480
Db 3571 TCCAGGATCTTGTGATTTGCGGCATTTCATCTCTAAGATCCGGCTTCCAAAC 3630
Qy 481 CTGCAATGCGCGCTGTACGAGAAATGCAAGATGCGTTACCGGGTTTACATTAAC 540
Db 3631 CTGCAATGCGCGCTGTACGAGAAATGCAAGATGCGTTACCGGGTTTACATTAAC 3690
Qy 541 CTTGAAAGGGCGGTGCATTTGACATGTAGTACTTATGTGTGAAACCTGATGCTGAAGGT 600
Db 3691 CTTGAAAGGGCGGTGCATTTGACATGTAGTACTTATGTGTGAAACCTGATGCTGAAGGT 3750
Qy 601 TCGGCAAGTTCTCTTCCAAACATCGACTTGTCTAGACCTGACAGACCGTTTGTGACCA 660
Db 3751 TCGGCAAGTTCTCTTCCAAACATCGACTTGTCTAGACCTGACAGACCGTTTGTGACCA 3810
Qy 661 TGTTCGATAGTGAAGATTCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 720
Db 3811 TGTTCGATAGTGAAGATTCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 3870
Qy 721 GTTCATGGCGCTGATTCGGGACTTCGGGACTTCGGGACTTCGGGACTTCGGGACTTC 780
Db 3871 ATCATGGCGCTGATTCGGGACTTCGGGACTTCGGGACTTCGGGACTTCGGGACTTC 3930
Qy 781 GACGATGTTACATATATGAAGATGATCGTGTGAGGCAAGCTTTGTCAATGCT 840
Db 3931 GACGATGTTACATATATGAAGATGATCGTGTGAGGCAAGCTTTGTCAATGCT 3990
Qy 841 TTCAGGAGCGCTCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 900
Db 3991 TTCAGGAGCGCTCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4050
Qy 901 TTCTGCTGTTTTTCTCTCTGAGCGTTACCGGCTGTTCTGATGAGCGCTTCCAAAT 960
Db 4051 TTCTGCTGTTTTTCTCTCTGAGCGTTACCGGCTGTTCTGATGAGCGCTTCCAAAT 4110
Qy 961 CCCGCGACGTGACACTTACTTGTCTACCAAGCGCTCCATACATGTGGAAGCCGGG 1020
Db 4111 CCCGCGACGTGACACTTACTTGTCTACCAAGCGCTCCATACATGTGGAAGCCGGG 4170
Qy 1021 CAGCTGCAACGAGCTTTCATCGCTTTCACAGGTTGAGGTCGTTCTTGAAGGAC 1080
Db 4171 CAGCTGCAACGAGCTTTCATCGCTTTCACAGGTTGAGGTCGTTCTTGAAGGAC 4230
Qy 1081 GGTTCATGAGCGAGATATGTGTGCTTCCGCTGTGCTATTAAGAGCGCTTGAA 1140
Db 4231 GGTTCATGAGCGAGATATGTGTGCTTCCGCTGTGCTATTAAGAGCGCTTGAA 4290
Qy 1141 TCAGGACATTAAGTGGGCTCATGACTCTGCAAAATTTGGTGAACGTTTGGGAGG 1200
Db 4291 TCAGGACATTAAGTGGGCTCATGACTCTGCAAAATTTGGTGAACGTTTGGGAGG 4350
Qy 1201 GATCCTCTCTTCAGGAGTAGAGAGATCTTTCGGGACACATCTCTGGGTGT -AA 1259

Db 4351 GAGTCCCTCTCTCAGCGATAGAGGATCTTTCTGGGACGATCTCTCGGTGTGAA 4410
 Qy 1260 ACATGAGTTTCTCTATGATGAGGACCTATTCAGCTATGAGGATGAGTCTGGCGG 1319
 Db 4411 ACATGAGTTTCTCTATGATGAGGACCTATTCAGCTATGAGGATGAGTCTGGCGG 1319
 Qy 1320 GTTGTGTCAGTTTGGAAAGCGGTTTATGAGATCTCTCGGTGTGATCAACGATTA 1379
 Db 4470 GTTGTGTCAGTTTGGAAAGCGGTTTATGAGATCTCTCGGTGTGATCAACGATTA 1379
 Qy 1380 TGAAGAAATCAGCGGATGTCCTTGAAGGATCTCAGAACTTCCAGCTGATCCATC 1439
 Db 4530 TGAAGAAATCAGCGGATGTCCTTGAAGGATCTCAGAACTTCCAGCTGATCCATC 1439
 Qy 1440 TGAAGGTTAAAGGATGTCCTGAGCGGATGTCCTGAGCTGATCCAGCTGATCCATC 1499
 Db 4590 TGAAGGTTAAAGGATGTCCTGAGCGGATGTCCTGAGCTGATCCAGCTGATCCATC 1499
 Qy 1500 TCAGAGGAAAGACAAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAATTTATGA 1559
 Db 4650 TCAGAGGAAAGACAAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAATTTATGA 1559
 Qy 1560 TAAAGTGTGTGATCTGAGACTCCGAAATATCCAACTCAGGATTCCTGATCCGAA 1619
 Db 4710 TAAAGTGTGTGATCTGAGACTCCGAAATATCCAACTCAGGATTCCTGATCCGAA 1619
 Qy 1620 TAAAGTGTGTGATCTGAGACTCCGAAATATCCAACTCAGGATTCCTGATCCGAA 1679
 Db 4770 TAAAGTGTGTGATCTGAGACTCCGAAATATCCAACTCAGGATTCCTGATCCGAA 1679
 Qy 1680 GTCAAAACTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739
 Db 4830 GTCAAAACTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739
 Qy 1740 TGTCTCTATGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1799
 Db 4890 TGTCTCTATGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1799
 Qy 1800 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1859
 Db 4950 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1859
 Qy 1860 GCGGCTCCCGGACAAAAGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1919
 Db 5010 GCGGCTCCCGGACAAAAGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1919
 Qy 1920 CCGGCGGTTTCCCGGACCACTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1979
 Db 5070 CCGGCGGTTTCCCGGACCACTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1979
 Qy 1980 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2039
 Db 5130 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2039
 Qy 2040 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2099
 Db 5190 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2099
 Qy 2100 TTAATGCGGCGGTTGAGTTTCTTCAAGGATGATGATGATGATGATGATGATGATGAT 2159
 Db 5250 TTAATGCGGCGGTTGAGTTTCTTCAAGGATGATGATGATGATGATGATGATGATGAT 2159
 Qy 2160 CCGGTTTAAAGCGGTTGAGTTTCTTCAAGGATGATGATGATGATGATGATGATGATGAT 2219
 Db 5310 CCGGTTTAAAGCGGTTGAGTTTCTTCAAGGATGATGATGATGATGATGATGATGATGAT 2219
 Qy 2220 TCTCTGCAAGCTCTTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2267
 Db 5370 TCTCTGCAAGCTCTTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2267

RESULT 8
 AE009419

LOCUS AE009419 10377 bp DNA linear BCT 20-DEC-2001
 DEFINITION Agrobacterium tumefaciens str. C58 Ti plasmid, section 2 of 21 of the complete sequence.
 ACCESSION AE009419 AE008690
 VERSION AE009419.1 GI:17744019
 KEYWORDS
 SOURCE
 ORGANISM
 Agrobacterium tumefaciens str. C58 (U. Washington)
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 1 (bases 1 to 10377)
 Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E., Zhou, Y., Woo, L., Kiteajima, J.P., Okura, V.K., Almeida Jr., N.F., Gillet, M., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I., Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V. and Nester, E.W.
 Direct Submission
 Submitted (27-SEP-2001) Department of Microbiology, University of Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA 98195-7242, USA
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 11743193
 2 (bases 1 to 10377)
 Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E., Zhou, Y., Woo, L., Kiteajima, J.P., Okura, V.K., Almeida Jr., N.F., Gillet, M., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I., Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V. and Nester, E.W.
 Direct Submission
 Submitted (27-SEP-2001) Department of Microbiology, University of Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA 98195-7242, USA
 FEATURES
 source
 1.10377
 /organism="Agrobacterium tumefaciens str. C58 (U. Washington)"
 /mol_type="genomic DNA"
 /strain="C58"
 /db_xref="taxon:180835"
 /plasmid="T1"
 319..513
 /gene="ipt"
 /note="synonym: Atu6006"
 319..513
 /gene="ipt"
 /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
 /codon_start=1
 /transl_table=11
 /product="isopentenyladenosine phosphotransferase"
 /protein_id="AA146242.1"
 /db_xref="GI:17744020"
 /translation="MDLYLIFGPTCTGKTAATLSEHTGIPVLAALDRICCPQLSTGSGRPAAYTAPOSIGORC"
 complement(380..1414)
 /gene="mas1"
 /note="synonym: Atu6007"
 complement(380..1414)
 /gene="mas1"
 /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
 /codon_start=1
 /transl_table=11
 /product="agrocinnopine synthesis reductase"

gene
/protein_id="AA146243.1"
/db_xref="GI:17744021"
/translation="MEGGGLPRKMFVENVYDCEKTELFSARVAGALTARRSNIILEV
SHGVLSVATLIGVELTAMORCNACVILHFKPPGSSNNVTRDSGVILVYAGSGIDK
AIDNLMNRGRFSLGARSVEKLEATGFGPNESEHFRFAEDLOTMEEVYATAMKE
GRIDLVNNAAGDGRVALAAGIDYEBLERWKMKNAGAPLMILCLPQLAKSSGRIIV
NIISMGQVRLNFEVGNMTHALSGTLTKAQHWGESSGVRTMEICPGVATQMSAMT
DEIDEGKIQPEHIAELVROAIRPNNAFVKIEVLICIQPLAVRSRSKVGSIYVGP
KRPGCCACAA"
complement(1622..2878)
/gene="acscx"
/note="synonym: Atu6008"
complement(1622..2878)
/gene="acscx"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="agrocinnopine synthase"
/protein_id="AA146244.1"
/db_xref="GI:17744022"
/translation="MEGTWPAPWTILHMDLPENYTPPEVLAHRSVPRILERLEEDLPQ
VIRKGFNLGKRIQECTASLLAAGQGRNISELDVCLTSDNVAIVSHDLTWVS
EKIGDKLFNEHSSKIKDVPVILREVSNGLIIOKYLETIDHILPTEIFSKYLIANS
ATIFLDGRNVEAHVIVAMLSHREYQVRVVLFTYEEYPHGAFDVAVINAQASAMR
KSIALMPALPEELCLARLROYTEPTVDLILAGKAMPMSLMODMRTIYAAHVPSG
VFNILIGOVVDKVLAFDSDOAAVLAYLYKEDTMRARPHLKRAVTRCYDPAL
LDSEGERSEIDIKTGRARHETDERKHIFMRKGTGNSATIDWVTSIDREDEMAIW
EMNQIDREVSLSPLDLNLTISK"
3636..4157
/gene="genes5"
/note="synonym: Atu6009"
3636..4157
/gene="genes5"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="gene 5 protein"
/protein_id="AA146245.1"
/db_xref="GI:17744023"
/translation="MNSIINTDVPIDPAIDEVKRCCEVACLPAGAPIPLIILNDLSL
TYVCSFOAMRKTAHRRFYDGVSDGVISTVPAEGITKEHMRGMHNNVCONTSNE
THLDAYIAFLPTSLQNPFSHMKIGCDSLAPSRVDFCEVLIANGKALFHNGLAKK
NPRCGGQMLPSC"
complement(4662..6065)
/gene="tms2"
/note="synonym: Atu6010"
complement(4662..6065)
/gene="tms2"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="indole acetamide hydrolase"
/protein_id="AA146246.1"
/db_xref="GI:17744024"
/translation="WVPTSLAQTLERLRKDYSCLELVETLIARCOAKPLNALIAT
DMWGLRSAAKIDRHGNAIGLGIPLCFPANATGIPTSAATPALIHLPITPSRV
AERLPGAGALPGASGMHHELSFGITSNNTAVGRNPNPBLPGSSGGLVAAVS
LMIGIGTGTGASVRLPALCGVVGFRPTARBRDRIIPVSPRTAGIAGCCVAD
IILDVYISGRSAKISPMPLKGLRIGLPTYFYVDLADVAFAAETIRILLIANGCTV
EADIPIHELSNGASLPALIEFPHALKYLDLDFVGTSESDVIKIRSPDVANITYA
QIDGHIISDEIELAOSFRPLQATYRNFRYLQDAILFTAPLAARAIGQESSVI
HNGSMNTEPKIYRVNDPSSNAGLPGLSLPACLTPLRPLVGEIDIGLAGSDHRLAIG
AALEKAINFSPFDATN"
6403..8670
/gene="tms1"
/note="synonym: Atu6011"
6403..8670
/gene="tms1"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"

gene
/codon_start=1
/transl_table=11
/product="tryptophan 2-monooxygenase"
/protein_id="AA146247.1"
/db_xref="GI:17744025"
/translation="MSASALDNOQDHFSTKRVNLIIMDKADELDRVSDAFSEAREAS
RGRITQIIGESGACIACRLADGRPEETISAEKVAALAYVIGKEIIGRIILIESPW
ARARVGLVAIADLAPCMDESEAOILQTLFISGKRCASDLSHPVAISIKIARSRT
LQMPREKGTTRKVTGFTLIEAVPDMVAAGRNLMKASGSPFTILLVYRSFP
DQCSBSGRIGFPPEVDYPRKRVAILIGISGLVYASLELHAGCDVITTYEASRVGCKL
WHAERKADPSVVAEWGMARFPAPASCLFFELRYGLSSMRPENPGTVDITNLVYQGLR
YMKAGQOPPKLFHRYVSGMRAFLKDFEGEDIVLASPYAITQALSGDIIRAHDSWQ
TWNLPGRGFSFSAIERIFLGTGHPGEGTMSFPHMDLFLKMGISGCGPYEESGFI
EILRLVINYEENORMCSGISELPRIRATOVNVGSVSORIRHVOVRAIIEKTKIK
ILKSGISRLYKVVYVTSGLANICLRHCITCOTTPRAVNOAVNDSHWTSKFL
TERKRWLDIIPSCVUMDGIKAVYGLDIEPODPNGKGLVLISTWEDSHKULAVPD
KKERFCLLDASISRSFPAPAOHLVPAQCAVDONVQHDMLTDBNAKFAFLNRGDF
VSEELFFQALDITNDTGVYLAGCSCSFTGQWVEGAIQTACNAVCAIHNCGILLADON
PLEHSWKRYNRYNRN"
9368..10090
/gene="lpc"
/note="synonym: Atu6012"
9368..10090
/gene="lpc"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="isopentenyl transferase"
/protein_id="AA146248.1"
/db_xref="GI:17744026"
/translation="MDLRLIFGPTGKTSTAVALAONOTGLPVLISDRVOCROSTG
SGRPVBEHKTSRLYLDPRPYKGIILAKQKHEHLMGVYVYEAHAGIILBEGSISL
LKQAKOSIYSADFRRHITRHELDABETDMNAKARVKQMLPAGALSTIIOELVDLWK
EERLRLIKEIDGYRYAMLPASQONITSDMLQLDADMDKLIHIGAOBYLIHAROE
QKPRVNAAYVGFEGHPGMY"
BASE COUNT 2745 a 2376 c 2438 g 2818 t
ORIGIN
Query Match 90.1%; Score 2044; DB 1; Length 10377;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2141; Conservative 0; Mismatches 125; Indels 2; Gaps 2;
QY 1 ATGTGAGTTTCACTCTCTGATATACGATCGATCACTCCCAACCAAAATGTGGAT 60
DB 6403 ATGTGAGTTTCACTCTCTGATATACGATCGATCACTCTCAACAAATGTGGAT 6462
QY 61 CTGACATGTCGATTAAGCGGATGAATTGACCCGAGGGTTCCGATGCTTCTAGAA 120
DB 6463 CTGATATGTCGATTAAGCGGATGAATTGACCCGAGGGTTCCGATGCTTCTAGAA 6522
QY 121 CGAAGAGCTTCAAGGGAAGAGATTAACAAGCTCCACGAGTGCAGCGCTGGGTTA 180
DB 6523 CGAAGAGCTTCAAGGGAAGAGATTAACAATCTCCGAGAGTGCAGCGCTGGGTTA 6582
QY 181 GCTTGAAGAGGCTGGCCATGCTGCTTCCCCGAGATCTCAGCTGCTGAAAAGTAA 240
DB 6583 GCTTGAAGAGGCTGGCCATGCTGCTTCCCCGAGATCTCAGCTGCTGAAAAGTAA 6642
QY 241 GTTCTCTCGCTTATATCATATTTGGCAAGAAATCTGGGGGGGATCTTGAATGAAA 300
DB 6643 GCTTCTCTCGCTTATATCATATTTGGCAAGAAATCTGGGGGGGATCTTGAATGAAA 6702
QY 301 CTTTGGCGCGGCAACAGTAGTGTCTGTTGCCATGCACTTGACACCAATTTGCATG 360
DB 6703 CTTTGGCGCGGCAACAGTAGTGTCTGTTGCCATGCACTTGACACCAATTTGCATG 6762
QY 361 GATTTCCTCGAGCAACATAATCCAGCCGCTTTTGGCTAGCGGTAAAGATGTGCA 420
DB 6763 GATTTCCTCGAGCAACATAATCCAGCCGCTTTTGGCTAGCGGTAAAGATGTGCA 6822
QY 421 CGATTGATCTAGCATTTCTGGGCATTTCAATCTTAAGTACGCGGCTTTGCAACC 480

Db 6823 TCCAGGATCTTATGATCTTGTGGCCATTTCAATCTTAAGAAGTCCGCGCTCCGAAACC 6882
 Qy 481 CTCGCAATGCGCGTGTACGAGAAATGGCAAGATGAAATGCGTTACCGGGTTTACCATTAACC 540
 Db 6883 CTGCAAAATGCCGCTTACGAGAAAGGACGACGAAAGCGCTTACCGGGTTTACCCTGACC 6342
 Qy 541 CTTGAAAGGGGCGCGTCCATTTGACATGTAAGCTTATGTTGGAACCTGATGCTGAAGGCT 600
 Db 6943 CTTGAAAGGCGCTTACCATTTGACATGTAAGCTTATGTTGGAACCTGATGCTGAAGGCT 7002
 Qy 601 TCGGCAAGGTTCTTTCCAAACATCGACTTGTGCTTATGACATGCTTTTGTGACCA 660
 Db 7003 TCGGCAAGGTTCTTTCCAAACATCGACTTGTGCTTATGACATGCTTTTGTGACCA 7062
 Qy 661 TGTTCGATGTGAGCGATCGGCTTCTTCCGGAAGATGTTCTTAAGCCGGAAGTGGCG 720
 Db 7063 TGTTCGATGTGAGCGATCGGCTTCTTCCGGAAGATGTTCTTAAGCCGGAAGTGGCG 7122
 Qy 721 GTTCATTGCGCTGCAATTTCCGGAAGTGTGCTGTTGCAACGAACTGCTTCAATGCTGGGTA 780
 Db 7123 ATCATTGGGCGTGGCATTTCCGGAAGTGTGCTGTTGCAACGAACTGCTTCAATGCTGGGTA 7182
 Qy 781 GAGCATGTTTCAATATATGAGCAAGTGTGTTGAGGCAAGCTTTGCTCAATGCT 840
 Db 7183 GACGATGTTTCAATATATGAGCAAGTGTGTTGAGGCAAGCTTTGCTCAATGCT 7242
 Qy 841 TTGAGGAGCGCTCTGATGTGTGCGGCAAAATGGGGGCGATGCGATTTCTCTCGCTGCA 900
 Db 7243 TTGAGGAGCGCTCTGATGTGTGCGGCAAAATGGGGGCGATGCGATTTCTCTCGCTGCA 7302
 Qy 901 TTCTGCTGTTTCTTCTCTCGAAGCTTACGCGCTGTCTTCAATGAGGCGTTCCCAAT 960
 Db 7303 TGTGTGCTGTTTCTTCTCTCGAAGCTTACGCGCTGTCTTCAATGAGGCGTTCCCAAT 7362
 Qy 961 CCGGCGCAAGTGCACACTTCTGTTACCAAGGCGTCCAAATCATGTGAAAGCGGG 1020
 Db 7363 CCGGCGCAAGTGCACACTTCTGTTACCAAGGCGTCCAAATCATGTGAAAGCGGG 7422
 Qy 1021 CAGCTGCAAGAGCTGTTCCATCGGCTTACCAAGGCTTGGGCGGCTTCTGAAGGAC 1080
 Db 7423 CAGCTGCAAGAGCTGTTCCATCGGCTTACCAAGGCTTGGGCGGCTTCTGAAGGAC 7482
 Qy 1081 GGTTCATGAGGAGATATGTTGCTTGGCTTGGCTGCTGCTATTAATCAAGCCTTGAAA 1140
 Db 7483 GGTTCATGAGGAGATATGTTGCTTGGCTTGGCTGCTGCTATTAATCAAGCCTTGAAA 7542
 Qy 1141 TCAGGACACTTATGTTGCTGCTATGACTCTGCGCAATTTGGCTGAACGCTTTGCGGAGG 1200
 Db 7543 TCAGGACACTTATGTTGCTGCTGCTATGACTCTGCGCAATTTGGCTGAACGCTTTGCGGAGG 7602
 Qy 1201 GAGTCTTCTCTTCAAGGATGAGAGGATCTTCTGGGCAACAATCTCTCGTGGTGT -AA 1259
 Db 7603 GAGTCTTCTCTTCAAGGATGAGAGGATCTTCTGGGCAACAATCTCTCGTGGTGT -AA 7662
 Qy 1260 ACATGAGTTTCTCTCATGATGAGGACTTATCAAGCTTAATGGGAATAGATCTGGCGG 1319
 Db 7663 ACATGAGTTTCTCTCATGATGAGGACTTATCAAGCTTAATGGGAATAGATCTGGCGG -GG 7721
 Qy 1320 GTTTGCTCAGTTTGAAGCGGCTTATTAAGATCTTCCGCTGCTGCTCAATCAAGGATA 1379
 Db 7722 GTTTGCTCAGTTTGAAGCGGCTTATTAAGATCTTCCGCTGCTGCTCAATCAAGGATA 7781
 Qy 1380 TGAAGAAATCAAGGATGAGGCTTCAAGGATCTCAAGCTTCAAGCTGAGTCCGATC 1439
 Db 7782 TGAAGAAATCAAGGATGAGGCTTCAAGGATCTCAAGCTTCAAGCTGAGTCCGATC 7841
 Qy 1440 TGAAGTGTAAACGCTGTCTGTGAGCAGCAGCAGTATGCAATGTTCAAGTCAAGGCGAT 1499
 Db 7842 TCAAGTGTAAACGCTGTCTGTGAGCAGCAGCAGTATGCAATGTTCAAGTCAAGGCGAT 7901
 Qy 1500 TCAAGGAAAGCAAAATTAAGATTAAGCTTAAAGCGGAGTATCTGAATTTATGA 1559
 Db 7902 TGAAGAAAGCAAAATTAAGATTAAGCTTAAAGCGGAGTATCTGAATTTATGA 7961

Qy 1560 TAAGTGTGTACATCTGACTGCAATATCCAACTCAGGCAATGCTGACATGCGA 1619
 Db 7962 TAAGTGTGTGTACATCTGACTGCAATATCCAACTCAGGCAATGCTGACATGCGA 8021
 Qy 1620 TACCAATATTTTCAAGGCAAGTGAACCAAGCGGTGTTAAACGCAATATGACAGATC 1679
 Db 8022 TACCAATATTTTCAAGGCAAGTGAACCAAGCGGTGTTAAACGCAATATGACAGATC 8081
 Qy 1680 GTCAAAACCTTCTCTGATGACTGAAGAAATTTCTGTTAGACATATCTCCGCTTG 1739
 Db 8082 GTCAAAACCTTCTCTGATGACTGAAGAAATTTCTGTTAGACATATCTCCGCTTG 8141
 Qy 1740 TGTCTCATGAGCAGGATGCAAAAGCAATGATGCTGCAATATGATGAGTCCGATCC 1799
 Db 8142 TGTCTCATGAGCAGGATGCAAAAGCAATGATGCTGCAATATGATGAGTCCGATCC 8201
 Qy 1800 GAATGTTAAAGCTTATGCTCATGATTAATATGAGGAGAGCACTCCCAAGCTGTT 1859
 Db 8202 GAATGTTAAAGCTTATGCTCATGATTAATATGAGGAGAGCACTCCCAAGCTGTT 8261
 Qy 1860 GGGGTCCTCCGCAAAAGAGGATTAATGCTGCTGCGGAGAGCAATTTGAGATCTTT 1919
 Db 8262 GGGGTCCTCCGCAAAAGAGGATTAATGCTGCTGCGGAGAGCAATTTGAGATCTTT 8321
 Qy 1920 CCGGCGGTTTGCAGCAAGCTTATTTCTGCTGCTGCTGTTAGACCAAAATGTTATCA 1979
 Db 8322 CCGGCGGTTTGCAGCAAGCTTATTTCTGCTGCTGCTGTTAGACCAAAATGTTATCA 8381
 Qy 1980 ACATGATGCTTACAGAGAAATCCGCGGAGGACTTTCAAACTCAACCGGCTGCTGA 2039
 Db 8382 ACATGATGCTTACAGAGAAATCCGCGGAGGACTTTCAAACTCAACCGGCTGCTGA 8441
 Qy 2040 GGAATTTTATTTGAGAACTTTCTTTCAAGCACTGACACGCGTAAATGATACCGGAGT 2099
 Db 8442 GGAATTTTATTTGAGAACTTTCTTTCAAGCGCTGAGACATACATATGATACCGGAGT 8501
 Qy 2100 TTACTTGGCGGCTTGGCAATGTTCTTCAAGTGTGATGGGTGAGGCTTATTCAGAC 2159
 Db 8502 TTACTTGGCGGCTTGGCAATGTTCTTCAAGTGTGATGGGTGAGGCGCTATTCAGAC 8561
 Qy 2160 CCGCTGTAACCGCGCTGCTGCAATTTCCCAATTTGTGAGGCAATTTGGCAAGGCGCA 2219
 Db 8562 CCGCTGTAACCGCGCTGCTGCAATTTCCCAATTTGTGAGGCAATTTGGCAAGGCGCA 8621
 Qy 2220 TCCTCTGGAACACTTGTGAGAGATTAATCAACGCAAGTAAATTA 2267
 Db 8622 TCCTCTGGAACACTTGTGAGAGATTAATCAACGCAAGTAAATTA 8669

RESULT 9
 AB016260
 LOCUS 206479 bp DNA circular BCT 12-JUN-2003
 DEFINITION Agrobacterium tumefaciens plasmid pTI-SAKURA, complete sequence.
 ACCESSION AB016260.1 AB007145 AB006857 AB006858
 VERSION AB016260.1 GI:6498173
 KEYWORDS
 SOURCE
 ORGANISM
 Agrobacterium tumefaciens (Rhizobium radiobacter)
 Agrobacterium tumefaciens
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.

REFERENCE
 1 Hattori, Y., Suzuki, K., Uraji, M., Ohta, N., Katoh, A. and Yoshida, K.
 Genome structure of pTI-SAKURA (II): strategy for DNA sequencing of
 a Japanese cherry-Ti plasmid
 Nucleic Acids Symp. Ser. 37, 159-160 (1997)
 MEDLINE 98247140
 PUBMED 9586048
 2 Hattori, Y., Suzuki, K., Ohta, N., Uraji, M., Katoh, A. and Yoshida, K.
 Genome structure of pTI-SAKURA (V): Complete nucleotide sequence of
 plasmid pTI-SAKURA's vir region in Agrobacterium tumefaciens
 Nucleic Acids Symp. Ser. 39, 265-266 (1998)
 JOURNAL


```

/transl_table=11
/protein_id="BAAB7631.1"
/db_xref="GI:6498179"
/translation="MRPCVPMVKAMTIOGROLEAPFAVWLTCGTSANLVRTTOPAI
SRIRLDEBIEIGISLEPRTGNRLPRTREACILFKEYSRHENGIOHIDRYAALBKXSH
GSLRVACYTAPALSTFMSVYQTEIADRPDVSVLDTVPSTQLELVSLDHDGISIL
AGSPQLTTEPVPSPFRAVCLLPGRHLEDEYVHATDLESELICISPVSLMQMDA
ADSCGVHKNRIESSLALNCDLVSRRMGVGIIDPFTADYSANPVIQRFDPVVPY
HFAIVPTDSEPPRLVSEFPRAALDLAKLPVETI"
complement(7308..8624)
/gene="clorf7"
/complement(7308..8624)
/gene="clorf8"
/function="Involved in conjugation"
/notes="Probable trbI gene"
/codon_start=1
/transl_table=11
/protein_id="BAA87632.1"
/db_xref="GI:6498180"
/translation="MIOSLKLGTAQOEELSGMRIRNLPILIIIVVLVVAFLVIFY
GLTRGLYPRGDEGPDITSGNPASTFADQIKRVDGIIIGEPOLVTLLOPTPYTOBE
KRVNPFPTQGRREARPAQJESQVWPAIRQROEPOYLREOROMARLQAND
AIDSPVANDKDFRAGADTAAMASATTVNSVSTATTSGADLYAALRAGLGQNDP
NGQSKDEPFENADLKEGLVPRVVPQSTVELKRSVLPATLITGINSDLPERITAO
VSONVYDSATGHRLLIPQGTLPGRYDSKVSFGQSRVLVMTDIIIFPNSSTLQIGWA
GTDAGVSGSRDVKRMLQTFSGAILIIVIGIDMAVPESSTLATODTADSAARN
FAETFRVADRTIORNDVQPTLEIRGVKFNVLVDQDIVPQTYR"
complement(8637..9113)
/gene="clorf8"
/complement(8637..9113)
/gene="clorf8"
/function="Involved in conjugation"
/notes="Probable trbH gene"

```

Query Match 90.1%; Score 2044; DB 1; Length 206479;
 Best Local Similarity 94.4%; Pred. No. 0;
 Matches 2141; Conservative 0; Mismatches 125; Indels 2; Gaps 2;

```

QY 1 ATGTGAGCTTACGCTCTTGTATTAACAGTGGCATCATGCCAACCAAAATGGTGGAT 60
DB 188587 ATGTGAGCTTACGCTCTTGTATTAACAGTGGCATCATGCCAACCAAAATGGTGGAT 188646
QY 61 CTGACAATGTCGATTAAGCGCGATGAATGACCGCGAGGTTTCCGATGCTTCTAGAA 120
DB 188647 CTGATATGTCGATTAAGCGCGATGAATGACCGCGAGGTTTCCGATGCTTCTAGAA 188706
QY 121 CGAAGAGCTTCTAGGCGAGAGGATTAATCTCAAGCTTCAACCGAGTGCAGCGCTGGTTA 180
DB 188707 CGTGAAGCTTCTAGGCGAGAGGATTAATCTCAAGCTTCAACCGAGTGCAGCGCTGGTTA 188766
QY 181 GCTTGCAGAAAGCGCGCGATGGTGGCTTCCCGAGATCTCAGTGGTGGAAAGGTAGCA 240
DB 188767 GCTTGCAGAAAGCGCGCGATGGTGGCTTCCCGAGATCTCAGTGGTGGAAAGGTAGCA 188826
QY 241 GTTCTCTCCGCTTATATCTAATTTGCAAGAAATTTGCGGCGCGATCTGATTCGAAA 300
DB 188827 GCTCTCTCCGCTTATATCTAATTTGCAAGAAATTTGCGGCGCGATCTGATTCGAAA 188886
QY 301 CTTTGGCGCGGCGACAGTGAAGTGTCTCTGTCGATCGACTTGAACACCAATTTGCATG 360
DB 188887 CTTTGGCGCGGCGACAGTGAAGTGTCTCTGTCGATCGACTTGAACACCAATTTGCATG 188946
QY 361 GATTTTCCGAGAGCAACTTAATCCAGCCCTGTTTTTGTGCGAGCGGTAAAGATGCGCA 420
DB 188947 GATTTTCCGAGAGCAACTTAATCCAGCCCTGTTTTTGTGCGAGCGGTAAAGATGCGCA 189006
QY 421 CCGATTGATCTTATGTCATTTCTGTGGCATTTCAATCTCTAAGATGCGCGCTTTGCAACC 480
DB 189007 TCCAGCATCTTATGTCATTTCTGTGGCATTTCAATCTCTAAGATGCGCGCTTTGCAACC 189066
QY 481 CTGCGCAATGCGCGCTGTACGAGAAATGCGACGATGAATGCGTTTACCGGAGTTTACCATACC 540
DB 189067 CTGCGCAATGCGCGCGCTGTACGAGAAATGCGACGATGAATGCGTTTACCGGAGTTTACCATACC 189126

```

```

QY 541 CTTGAAGGCGCGTGCATTTTGACATGATGCTTAATGTCGAAAACCTGATGCTGAAGGT 600
DB 189127 CTTGAAGGCGCGTGCATTTTGACATGATGCTTAATGTCGAAAACCTGATGCTGAAGGT 189186
QY 601 TCGGCAAGTTCCTTTCCAAACATTCAGCTTCTCTACAGATCGACACCGTTTTTGAACCA 660
DB 189187 TCGGCAAGTTCCTTTCCAAACATTCAGCTTCTCTACAGATCGACACCGTTTTTGAACCA 189246
QY 661 TGTTCGATAGTGAAGCAATCGGCTCTTCCGAGAGATGTTCTTAAGCGCAAGTGGCG 720
DB 189247 TGTTCGATAGTGAAGCAATCGGCTCTTCCGAGAGATGTTCTTAAGCGCAAGTGGCG 189306
QY 721 GTCATTTGGCGTGCATTTCCGAGCTCGTGTGCGCAACGAACTGCTTCACTGCGAGTA 780
DB 189307 ATCATTTGGCGTGCATTTCCGAGCTCGTGTGCGCAACGAACTGCTTCACTGCGAGTA 189366
QY 781 GACGATGTTACAAATATATGAGCAAGTGAATCGTGTGAGGCAAGCTTTGTCACATGCT 840
DB 189367 GACGATGTTACAAATATATGAGCAAGTGAATCGGTTGAGGCAAGCTTTGTCACATGCT 189426
QY 841 TTCAGGAGCGCTCTGATGTCGTCGCGCAAAATGGGGCGCATGCGATTTCTCTGCTGCA 900
DB 189427 TTCAGGAGCGCTCTGATGTCGTCGCGCAAAATGGGGCGCATGCGATTTCTCTGCTGCA 189486
QY 901 TTCGCTGTTGTTTTCTTCTCGACGCTTACGCGCTGCTTCAATGAGCCGTTCCCAAT 960
DB 189487 TTCGCTGTTGTTTTCTTCTCGACGCGTACGCGCTGCTTCAATGAGCCGTTCCCAAT 189546
QY 961 CCCGCGACAGTGCACCTTACTGTTCTACCAAGCGCTCCAAATACATGTTGAAAACCGCG 1020
DB 189547 CCCGCGACAGTGCACCTTACTGTTCTACCAAGCGCTCCAAATACATGTTGAAAACCGCG 189606
QY 1021 CAGCTGCACCGAGCGTTCATTCGCGTTTACAGCGTTGGCGTGGTCTTGAAGAC 1080
DB 189607 CAGCTGCACCGAGCGTTCATTCGCGTTTACAGCGTTGGCGTGGTCTTGAAGAC 189666
QY 1081 GCTTTTCATGAGCGAGATATTTGTTGGCTTCCGCTGCTGCTTATTACTCAGGCTTGA 1140
DB 189667 GCTTTTCATGAGCGAGATATTTGTTGGCTTCCGCTGCTGCTTATTACTCAGGCTTGA 189726
QY 1141 TCAGAACACTTAAGTGGGCTCAATGCTCTGCGCAAAATTTGGCTGAACCGTTTCGGAGG 1200
DB 189727 TCAGAACACTTAAGTGGGCTCAATGCTCTGCGCAAAATTTGGCTGAACCGTTTCGGAGG 189786
QY 1201 GAGTCTTCTCTCAGGATAGAGAGATCTTCTGCGCAACATCTCTGCTGGT- AA 1259
DB 189787 GAGTCTTCTCTCAGGATAGAGAGATCTTCTGCGCAACATCTCTGCTGGTGGT- AA 189846
QY 1260 ACATGAGATTTTCTCATGATTTGGAACCTATTCAGCTAATGGGAAATGAGATCTGGCGG 1319
DB 189847 ACATGAGATTTTCTCATGATTTGGAACCTATTCAGCTAATGGGAAATGAGATCTGGCGG 189906
QY 1320 GTTTGCTCAGTTTTTGAAGCGGTTTATGAGATCTCTGCTGTTGTCATCAACGATA 1379
DB 189906 GTTTGCTCAGTTTTTGAAGCGGTTTATGAGATCTCTGCTGTTGTCATCAACGATA 189965
QY 1380 TGAAGAAAATCAACGAGATGGCCCTGAAGGAATCTCAGAACTTCCAGTGGATGCGATC 1439
DB 189966 TGAAGAAAATCAACGAGATGGCCCTGAAGGAATCTCAGAACTTCCAGTGGATGCGATC 190025
QY 1440 TGAAGTGTAAACGATGTCCTGTGAGCCAGCCATATGCCATGTTCAAGTCAAGCGGAT 1499
DB 190026 TGAAGTGTAAACGATGTCCTGTGAGCCAGCCATATGCCATGTTCAAGTCAAGCGGAT 190085
QY 1500 TCAAGAGAAAGACAAATAAAGATTAAGAGTGAAGCGGAGATATCTGAATCTTAATGA 1559
DB 190086 TCAAGAGAGAAAGACAAATAAAGATTAAGAGTGAAGCGGAGATATCTGAATCTTAATGA 190145
QY 1560 TAAAGTGTGTCAATCTGGAATTCGCAATTAATCAACTGAGGATGCTGACATGCGA 1619
DB 190146 TAAAGTGTGTGTCAATCTGGAATTCGCAATTAATCAACTGAGGATGCTGACATGCGA 190205

```

Qy	1620	TACCAATATTTTTCAGCACCAGTGAACCAAGCGGTGATATACAGCCATATGACAGAGATC	1679
Db	190206	TACACACATTTTTCGTGSCACAGTGAACCAAGCGGTGATATACAGCCATATGACAGGCTC	190265
Qy	1680	GTCAAACTCTTCTCTGATGACTGAACGAAAATCTGGTTAGACCATATCTCCCTCTTG	1739
Db	190266	GTCAAACTCTTCTCTGCTGACTGAACGAAAATTTTGGTTAGACCATATCTCCCTCTTG	190325
Qy	1740	TGTCCTATGAGAGGAGATGCGCAAAAGAGTATTTGCTGAGCATATAGTCGACAGATCC	1799
Db	190326	TGTCCTATGAGAGGAGATGCGCAAAAGAGTATTTGCTGAGCATATAGTCGACAGATCC	190385
Qy	1800	GAATGTAAAGTCTAGTGTCTCATAGTTATACATGGAGGAGACACTCCCAAGCTGT	1859
Db	190386	GAATGTAAAGTCTAGTGTCTCATAGTTATACATGGAGGAGACACTCCCAAGCTGT	190445
Qy	1860	GGGGGTCCCCGACAAAAGAGCGATTATGTCTGCTGGGAGCGCAATTTTGAAGATCTTT	1919
Db	190446	GGGGGTCCCCGACAAAAGAGCGATTATGTCTGCTGGGAGCGCAATTTTGAAGATCTTT	190505
Qy	1920	CCGGGGCTTGGCCAGACCTATTTCTGCTGGCGCTGATTAAGACCAAAATGTTATCA	1979
Db	190506	CCGGGGCTTGGCCAGACCTATTTCTGCTGGCGCTGATTAAGACCAAAATGTTATCA	190565
Qy	1980	ACATGATTTGGCTTACAGACGAGATGCGGGGGAGCTTTCAAACTCAACCGCGTGTGA	2039
Db	190566	ACATGATTTGGCTTACAGACGAGATGCGGGGGAGCTTTCAAACTCAACCGCGTGTGA	190625
Qy	2040	GGAATTTTATTTCTGAAGAACTTTTCTTCAAGACCTGAGACCGGCTAATGATACCGGAGT	2099
Db	190626	GGAATTTTATTTCTGAAGAACTTTTCTTCAAGCGGCTGACATGATGATACCGGAGT	190685
Qy	2100	TTTACTTGGCGGGTTCAGTGTCTCTTCAACAGTGGATGGGTGAGGGTCTATTCAGAC	2159
Db	190686	TTTACTTGGCGGGTTCAGTGTCTCTTCAACCGGTGATGGGTGAGGGTCTATTCAGAC	190745
Qy	2160	CGGCTGTAAGCGCGTCTGTGCAATTAATCAACAATGTGAGGACATTTTGGCAAGGGCAA	2219
Db	190746	CGGCTGTAAGCGCGTCTGTGCAATTAATCAACAATGTGAGGAGTATTTTGGCAAGGACAA	190805
Qy	2220	TCCTCTCGAAGCACTCTTGAAGAGATTAACACCGCAGTAGAATTA	2267
Db	190806	TCCTCTCGAAGCACTCTTGAAGAGATTAACACCGCAGTAGAATTA	190853
RESULT 10			
AF126446			
LOCUS	5968 bp	DNA	linear BCT 25-FEB-2000
DEFINITION	Agrobacterium tumefaciens plasmid pTiC58 agrocinopline synthase (acs) gene, partial cds; and 5 protein (5), indole-3-acetamide hydrolase (iaah), and tryptophan monoxygenase (iaam) genes, complete cds.		
ACCESSION	AF126446		
VERSION	AF126446.1	GI:4836565	
KEYWORDS	Agrobacterium tumefaciens (rhizobium radiobacter)		
SOURCE	Agrobacterium tumefaciens; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.		
REFERENCE	1 (bases 1 to 5968)		
AUTHORS	Osten, L., Salomone, J. Y., Helfer, A., Schmidt, J., Hammann, P. and De Ruffray, P.		
TITLE	Sequence and functional analysis of the left-hand part of the T-region from the nopaline-type Ti plasmid, pTiC58		
JOURNAL	Plant Mol. Biol. 41 (6), 765-776 (1999)		
MEDLINE	20199402		
PUBMED	10737141		
REFERENCE	2 (bases 1 to 5968)		
AUTHORS	Osten, L., De Ruffray, P. and Hammann, P.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-FEB-1999) Cell Biology, IBM, rue du General Zimmer		
FEATURES	12, Strasbourg 67804, France		
	Location/Qualifiers		

source	1. .5968
	/organism="Agrobacterium tumefaciens"
	/mol_type="genomic DNA"
	/strain="C58"
	/db_xref="taxon:358"
	/plasmid="pTiC58"
	/note="Ti plasmid; T-DNA"
gene	complement(<1. .19)
	/gene="acs"
	complement(<1. .19)
CDS	/gene="acs"
	/codon_start=1
	/product="agrocinopline synthase"
	/protein_id="AAD30486.1"
	/db_xref="GI:4836566"
	/translation="MFNLGK"
gene	782. .1462
	/gene="5"
CDS	782. .1462
	/gene="5"
	/function="may catalyze synthesis of indole-3-lactate"
	/codon_start=1
	/product="5 proleuin"
	/protein_id="AAD30487.1"
	/db_xref="GI:4836567"
	/translation="MYHSRPIFNIIIDSNIOREIKVLRLHTEIAYRSFAOEDLIPA
	ORSMWNSINTDVPIDPAIDEVKRCFVACLPAGIPLNTIINDSLTYVCSFGAM
	RKYAHKRFYDGVSDGVISTVPAEIGTETKMSMNNVCONTSNETHDIDAYAF
	LPTSLONPFSHMKIGCSFLAPSRVDPCEIIVAGKALFHDNGKKKPKWMSAP
	TFILKRLV"
	complement(1953. .3353)
gene	/gene="iaah"
CDS	complement(1953. .3353)
	/gene="iaah"
	/function="synthesizes indole-3-acetic acid from
	indole-3-acetamide"
	/codon_start=1
	/product="indole-3-acetamide hydrolase"
	/protein_id="AAD30488.1"
	/db_xref="GI:4836568"
	/translation="WVPIRSLAQTLERLRKRYSCLEELVKTIARCOAKPPINLLAT
	DWDGLRTRAKTIDRHGNGGLGGLPLCKGNATGIFTSATPAILHLPNIBSRV
	PQTLFSAQPLGASGMHLSFGITSNVATGAVRNPVPSLIPGSSGVAVAASR
	LMAGIGDTGASVRLPALCGVGFRRPTIARPPDRILIPVPTDITAGIIACVADV
	ILDDVIGSRSAKILPMLKGRIGLPTTFYDDLDADVAFAETIRLLANRGVTFV
	EDIDHLEINSGASLPTALVEFPAKLLYDDPGTFLTFRRIIRSPDVANIVSAO
	IDGHOISNDREYLAQSPRPRIQATYRFRLYQDDAILFPAPLAARICQESSYIH
	NQSMWTFKIVRYNADPPSSNAGLPGLSPALCTPRLPVGMEIDGLAGDHRLLAIGA
	ALEKAINPSPFPDAFN"
	3690. .5957
gene	/gene="iaam"
CDS	3690. .5957
	/gene="iaam"
	/function="synthesizes indole-3-acetamide from tryptophan"
	/codon_start=1
	/transl_table=11
	/product="tryptophan monoxygenase"
	/protein_id="AAD30489.1"
	/db_xref="GI:4836569"
	/translation="WSASLTDNQCDSHTKXVDLIMDKADELDRVSDAFSREAS
	RNRMTIOISGCSAGIACRLADGRFPEISABEKVAALSAVYVGEIIRLIESBPW
	AARVSGVAIDLAFPCMDPSEAOQLQTLFLSGRCASSDSHFVAISISTASRT
	LQMPYBKGTTRKRVGFTLTLEAVPDMVAAGRLMLKASGSPPTIDLIDYRSFF
	DOCSDSGRIPPEPDPKPKVAILGAGISGLVVAELIHAAGVDVLIVASPRVGLK
	WSHAKDAPSVIAEKARFPFPAASCLFFELRYGLSSRRPFPNGETVDITNLYQGLR
	TMWKGQDPKLFHRVYSGWRKFLKDGPHBGDIVLASPVATQALKSGDIRAHDSWQ
	TWLNRFGRSFSFSAERILFLGTHPPGGETWSPHMDLFLKMGISGSGGVPVESGFI
	ELIRLVINGEYENORMGSEGISLPRRIATQVNVDSVQIRHYQVRAIIEKEXTIK
	IRKSGISELYKVVVTSGLANIIOGRHCTCTTTFRAVNDVAVNSHWSGSKFLIL
	TERKFWLHIILPSCVLMDGIAKAVYCLDYEPDPPKNGKGLVILSYTWEDDSHLLAVPD

KKERFCLLRDAISRPFAQHLPACADYDQNVODMLTDENAGAFKLRNEDF
YSEHFEQALDNDTGVLYAGSCSFTGWVBGAJQTAQNAVCAIHHNGGTLANDN
PLEHSMKRYNRYRNRN"
BASE COUNT 1543 a 1333 c 1472 g 1620 t
ORIGIN

Query Match 90.1%; Score 2042.4; DB 1; Length 5968;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2140; Conservative 0; Mismatches 126; Indels 2; Gaps 2;

QY 1 ATGTGAGCTTCACTCTCTTGTATACAGTGCATCATCTCCACCAAAATGTGGAT 60
DB ATGTGAGCTTCACTCTCTTGTATACAGTGCATCATCTCTCCAAATGTGGAT 3749
QY 61 CTGACATGCTCATTAAGGGGATGATTTGACCGGATTTCCGATGCTTTAGAA 120
DB CTGATATGCTCATTAAGGCTGATGATGACCGGATTTCCGATGCTTTAGAA 3809
QY 121 CGAAGAGCTTCTAGGGGAGAGGATTTACTCAACGCTCAACGAGTCAAGCTGGGTTA 180
DB CGTGAAGCTTCTAGGGGAGAGGATTTACTCAATCTCCGGCGAGTCAAGCTGGGTTA 3869
QY 181 GCTTGCAGAAAGCTGGCCGATGCTGCTTCCCGAGATCTCACTGCTGAGAAAGTAGCA 240
DB GCTTGCAGAAAGCTGGCCGATGCTGCTTCCCGAGATCTCACTGCTGAGAAAGTAGCA 3929
QY 241 GTTCTCTCCGCTTATATCTATTTGGCAAGAAATTTGGGGGGGATTAAGTGAATGAAA 300
DB GCCCTCTCCGCTTATATCTATTTGGCAAGAAATTTGGGGGGGATTAAGTGAATGAAA 3989
QY 301 CTTTGGGCGGGGCAAGTGAAGTGTCTGTTGCCATTCGACTTGAACACCATTTGCATG 360
DB CTTTGGGCGGGGCAAGTGAAGTGTCTGTTGCCATTCGACTTGAACACCATTTGTATG 4049
QY 361 GATTTCCTCGAAGCAACAATTCGAACCCCTGTTTGTGAGCGGTAAGTAGTGA 420
DB GATTTCCTCGAAGCAACAATTCGAACCCCTGTTTGTGAGCGGTAAGTAGTGA 4109
QY 421 CCGATTTGATTTAGTCAATTTCTGGCCATTTCAATCTCTAAGACTGGCGGCTTTGCAAC 480
DB TCAGCATTTGATTTAGTCAATTTCTGGCCATTTCAATCTCTAAGACTGGCGGCTTTGCAAC 4169
QY 481 CTGCAATGCGCGCTGTAGCAAGATGCGATGAATGCGTTACCGGGTTTACATTAAC 540
DB CTGCAATGCGCGCTGTAGCAAGATGCGATGAATGCGTTACCGGGTTTACATTAAC 4229
QY 541 CTTGAAGGGGCGCTGCAATTTGACATGTAAGTGTGAAACCTGATGCTGAAGGCT 600
DB CTTGAAGGGGCGCTGCAATTTGACATGTAAGTGTGAAACCTGATGCTGAAGGCT 4289
QY 601 TCGGAGGCTTCTTCCAAATTCGATTTCTAGCACTGCAACCGGTTTGTGACAA 660
DB TCGGAGGCTTCTTCCAAATTCGATTTCTAGCACTGCAACCGGTTTGTGACAA 4349
QY 661 TGTTCGATAGTGAAGGATGCGCTTCTTCCGAGATGTTCTTAAGCCGAAAGTGGC 720
DB TGTTCGATAGTGAAGGATGCGCTTCTTCCGAGATGTTCTTAAGCCGAAAGTGGC 4409
QY 721 GTCATTTGGGCTGCAATTTCCGACTGCTGCTGCAACGAACTGCTTCAATGCTGGGTTA 780
DB ATCATTTGGGCTGCAATTTCCGACTGCTGCTGCAACGAACTGCTTCAATGCTGGGTTA 4469
QY 781 GAGATGTTACATATATGAAGCAAGTATGCTTTGAGAGCAAGCTTTGTCAATGCT 840
DB GAGATGTTACATATATGAAGCAAGTATGCTTTGAGAGCAAGCTTTGTCAATGCT 4529
QY 841 TTCAGGAGCCTCTAGTGTGCTGCTGCAAGTGGGGGAGTTCGATTTCTCCGCTGCA 900
DB TTCAGGAGCCTCTAGTGTGCTGCTGCAAGTGGGGGAGTTCGATTTCTCCGCTGCA 4589
QY 901 TTTGCTTTGTTTCTTCTCTGAGCGTGAAGCGCTGTCTTCTGATGAGCGGTTCCCAAT 960
DB TTTGCTTTGTTTCTTCTCTGAGCGTGAAGCGCTGTCTTCTGATGAGCGGTTCCCAAT 4649

QY 961 CCCGGCAGTGCACACTTACTTGTCTACCAAGGCGCTCAATACATGTGAAAGCGGG 1020
DB CCCGGCAGTGCACACTTACTTGTCTACCAAGGCGCTCCGATACATGTGAAAGCGGG 4709
QY 1021 CAGCTGCACCGAAGCTGTTCCATCGCGTTTACCAAGGTTGGGCGGTTCTTGAAGAC 1080
DB CAGCTGCACCGAAGCTGTTCCATCGCGTTTACCAAGGTTGGGCGGTTCTTGAAGAC 4769
QY 1081 GGTTCATGAGGAGATATGTGTGGCTTGGCTGTGGCTATTAAGGCTTGAAG 1140
DB GGTTCATGAGGAGATATGTGTGGCTTGGCTGTGGCTATTAAGGCTTGAAG 4829
QY 1141 TCAGGACATTTAGGGGCTCATGACTCTGGCAATTTGGCTGAACGTTTCGGAGG 1200
DB TCAGGACATTTAGGGGCTCATGACTCTGGCAATTTGGCTGAACGTTTCGGAGG 4889
QY 1201 GAGTCTTCTCTTCAAGGATAGAGAGATCTTTCTGGGACACATCTCTGTGTGT-AA 1259
DB GAGTCTTCTCTTCAAGGATAGAGAGATCTTTCTGGGACACATCTCTGTGTGTAA 4949
QY 1260 ACATGAGTTTCTCATGATTTGGGACCTATTCAGCTAATGGGAATAGATCTGGCGG 1319
DB ACATGAGTTTCTCATGATTTGGGACCTATTCAGCTAATGGGAATAGATCTGGC-GG 5008
QY 1320 GTTGGTCCAGTTTGTGAAGGGGTTTATTTAGATCTCCCGCTTGTGATCAACGGATA 1379
DB GTTGGTCCAGTTTGTGAAGGGGTTTATTTAGATCTCCCGCTTGTGATCAACGGATA 5068
QY 1380 TGAAGAAATTCAGCGGATGTCCCTGAAGGAATCTCAGAACTTCAAGTGTGCATC 1439
DB TGAAGAAATTCAGCGGATGTCTCTGAAGGAATCTCAGAACTTCAAGTGTGCATC 5128
QY 1440 TGAAGTGTTAAGGTGTGTGTGTGAGCCGCAATATGCCATTTCAACTCAGGCGAT 1499
DB TGAAGTGTTAAGGTGTGTGTGTGAGCCGCAATATGCCATTTCAACTCAGGCGAT 5188
QY 1500 TCAGAGGAAAGACAAAATTAAGATTAAGGCTTAAGAGGGGATATCTGAACCTTTATGA 1559
DB TCAGAGGAAAGACAAAATTAAGATTAAGGCTTAAGAGGGGATATCTGAACCTTTATGA 5248
QY 1560 TAAAGTGTGTGCATCTGCACTGCAAAATATCAACTCAGGCAATTCCTGACATGCGA 1619
DB TAAAGTGTGTGCATCTGCACTGCAAAATATCAACTCAGGCAATTCCTGACATGCGA 5308
QY 1620 TACCAATTTTTTCAAGCACAGTGAACCAAGCGGTTGATTAAGCCATATGACAGATC 1679
DB TACCAATTTTTTCAAGCACAGTGAACCAAGCGGTTGATTAAGCCATATGACAGATC 5368
QY 1680 GTCAAAACCTTCTGATGACTGAACGAAATCTGCTTGAACCATATCTCCCGCTTG 1739
DB GTCAAAACCTTCTGATGACTGAACGAAATCTGCTTGAACCATATCTCCCGCTTG 5369
QY 1740 TGTCTCATGAGAGGATGCAAAAGCAAGTATGCTCTGAGCTATGATGCGAGATCC 1799
DB TGTCTCATGAGAGGATGCAAAAGCAAGTATGCTCTGAGCTATGATGCGAGATCC 5429
QY 1800 GAATGTTAAAGCTGTGTGTCTCATGATTATATGAGGAGAGACTCCCAAGCGTT 1859
DB GAATGTTAAAGCTGTGTGTCTCATGATTATATGAGGAGAGACTCCCAAGCGTT 5548
QY 1860 GGGGATCCCGCAAAAAGAGGATTAATGCTGCTGGGAGCGCAATTTGAGATCTTT 1919
DB GGGGATCCCGCAAAAAGAGGATTAATGCTGCTGGGAGCGCAATTTGAGATCTTT 5549
QY 1920 CCCGGGTTTCCAGACACTATTTCTGCTGCGCTGATTAAGACCAAAATGTTATCA .1979
DB CCCGGGTTTCCAGACACTATTTCTGCTGCGCTGATTAAGACCAAAATGTTATCA 5608
QY 1980 ACATGATTTGCTTACAGACGAGATCCGGGGAGCTTTCAAACTCAACCGGCTGTGTA 2039
DB ACATGATTTGCTTACAGACGAGATCCGGGGAGCTTTCAAACTCAACCGGCTGTGTA 5669
QY 5669 ACATGATTTGCTTACAGACGAGATCCGGGGAGCTTTCAAACTCAACCGGCTGTGTA 5728

Oy	2040	GGATTTTATCTGAAGAACTTTCTTCACGACATGCAACGGCTAATGATTACCGAGT	2099
Db	5729	GGATTTTATCTGAAGAACTTTCTTCACGCCGTGCACATGACTTAATGATCCGGAGT	5786
Oy	2100	TTACTTGCGGGGTGCAGTTGTTCCTTCACAAGTGGATGGTGGAGGGTGCTATTACAG	2158
Db	5789	TTACTTGCGGGGTGCAGTTGTTCCTTCACCGGTGATGGGTGGAGGGCGCTATTACAGC	5848
Oy	2160	CGCGGTAAAGCCGCTCTGTGCATTATACCAATTGTGGAGGCATTTTGGCAAAGGC	2219
Db	5849	CGCGGTAAAGCCGCTCTGTGCATTATACCAAAATTGTGGAGGTATTTTGGCAAAGCA	5908
Oy	2220	TCCCTCGAACACTCTTGGAGAAGATATACTAACCGCAGTAGAAATTA	2267
Db	5909	TCCCTCGAACACTCTTGGAGAAGATATACTAACCGCATAGAAATTA	5956
RESULT 11			
AB025110			
LOCUS			
DEFINITION	AB025110	4660 bp DNA linear BCT 27-MAR-1999	
ACCSSION	AB025110	Agrobacterium tumefaciens gene for indole acetamide hydrolase and tryptophan monooxygenase, complete cds.	
VERSION	AB025110.1	GI:4586311	
KEYWORDS	tryptophan monooxygenase; indole acetamide hydrolase.		
SOURCE	Agrobacterium tumefaciens (Rhizobium radiobacter)		
ORGANISM	Agrobacterium tumefaciens		
REFERENCE	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.		
AUTHORS	1 (bases 1 to 4660) Kasahara,T., Endo,S., Ebinuma,H., Sugita,K., Kawaoka,A. and Moritame,N.		
JOURNAL	Direct Submission		
AUTHORS	Submitted (18-MAR-1999) Takehide Kasahara, Nippon Paper Industries Co., Ltd. R&D Div., Wood Bio-engineering Central Research Lab., 5-21-1, Oji, Kita-ku, Tokyo 114-0002, Japan (E-mail:take@ntc.biglobe.ne.jp, Tel:81-3-3911-5106(ex.292), Fax:81-3-3914-3350)		
FEATURES	Location/Qualifiers		
source	1..4660		
gene	/organism="Agrobacterium tumefaciens"		
CDS	/mol_type="genomic DNA" /strain="PO2" /db_xref="taxon:358" complement(353..1756) /gene="iaah" complement(353..1756) /gene="iaah" /codon_start=1 /transl_table=1 /product="indole acetamide hydrolase" /protein_id="BAA76345.1" /db_xref="GI:4586312"		
gene	/translation="MVPITSLAQTLERLRKRDYSCLEVETLIRQCAAKLNALLLATDWDGRSAKKNDHRGNAGLDLCGLPCEFKNIATGVPEPSAAPPALINHLPKIPSRRAERLFSGALPGASGNHIELFSGITSNNYAVGANPNPPLIPGSSGVVAVASRPAFMLGIGTDGASVALPALCGVVGFRTPLGRVPRDRILPPSPTRDTGIIAOCVADVDIILIDHOISGRSAKISPMLKGLRIGLPTTYDDDAVAPAEFTTRLRLANGTFVFAEADIPIVELNSGASLPALVEFPALKKYVDLVGTVGPSVDVKGISPVNIVSANOIHDOJNSDEYLAROSFRFRLOATYRNTRYILODAIIEPTPLPAKAIGQSSSVITHNSNMWTFKIYKRVNDPSSNAGLPGSLPACTLPDRUPVGMEIDGLAGSHRLLAIGAALKEAIINFSSFPAPFN"		
gene	2096..4363 /gene="iaam" 2096..4363 /gene="iaam" /codon_start=1		
CDS			

[illegible]

OY	72	GTCAATGGCCGTGCATTTCCGACATCGTGGTGGCAACGAACGTCTTCAATCGTGGGTA	780
Db	2816	ATCATTTGGCGGTGGCATTTCCGAGCTCGTGTAGCAAGCGAATCGTTCAATGCTGGTGA	2875
OY	781	GACATGTTACAAATATATATGAAGCAAGTATGCTGTGGAGGCAAGCTTTGGTCAATGCT	840
Db	2876	GACATGTTACAAATATATGAAGCAAGTATGCGGTTGGAGGCAAGCTTTGGTCAATGCT	2935
OY	841	TTCAAGGACGCTCCTAGTGTGTGGCCGAAATGGGGGCGATGCGATTTCTCTGTGCA	900
Db	2936	TTCAAGGATGCTCCACAGCGTGTGGCCGAAATGGGGGCGATGCGATTTCTCTGTGCA	2995
OY	901	TTCTGCTTTGTTTTTCTTCTCGACGCTTACCGGCTGTCTTTCGATGAGGCGCTTCCCAAT	960
Db	2996	TCGTGCTGTTTTTCTTCTCGACGCGTACCGGCTGTCTTTCGATGAGGCGCTTCCCAAT	3055
OY	961	CCCGGCAAGTCGACACTTACTTGGGTCACCAAGGCGTCCAAATCATGTGAAAGCCGGG	1020
Db	3056	CCCGGCAAGTCGACACTTACTTGGTCTTACCAAGGCTTCGATGATGTGAAAGCCGGG	3115
OY	1021	CAGCTGCCACGGAAGCTGTTCATCGCTTTACAACGGTTGGCGTTCGTTTGAAGAC	1080
Db	3116	CAGAGCCACCGAAGCTTCTCATCGCTTTACAGCGGTTGGCGTTCGTTTGAAGGAC	3175
OY	1081	GCTTTTCAATGAGCGAGATATTGTGTGGCTTCGCTGTGCTATTTACTCAGGCTTGA	1140
Db	3176	GCTTTTCAATGAGGAGATATTGTGTGGCTTCGCTGTGTTATTTACTCAAGCCTTGA	3235
OY	1141	TCAGGACATTTAGGTTGGGCTCATGACTCGTGGCAAAATTTGGGCTGAACCGTTTGGGAG	1200
Db	3236	TCAGGAGCATTTAGGCGGGCTCATGACTCTGGCAAACTTGGCTGAACCGTTTGGGAG	3295
OY	1201	GAGTCTCTTCTTCAAGGATAGAGAGATCTTTCTGGGACACATCTCTGTGTGT-AA	1259
Db	3296	GAGTCTCTTCTTCAAGGATAGAGAGATCTTTCTGGGACGACATCTCTGTGTGTGA	3355
OY	1260	ACATGGAATTTTCTCATGATTTGGGACCTATTCAAGCTAATGGGATATGGATCTGGCGG	1319
Db	3356	ACATGGAATTTTCTCATGATTTGGGACCTATTCAAGCTAATGGGATATGGATCTGGCGG	3414
OY	1320	GTTTGGTCCAGTTTTTGAAGCGGGTTAATTGAGATCCTCGGCTTGGTCATCAACGATA	1379
Db	3415	GTTTGGTCCAGTTTTTGAAGCGGGTTAATTGAGATCCTTGGCTTGGTCATTAACGATA	3474
OY	1380	TGAAGAAATCAGCGGATGTGCCCTGAAGGAATCTCAGAACCTTCCACGTCGGATCGCATC	1439
Db	3475	TGAAGAAATCAGCGGATGTGTCTGTGAAGGAATCTCAGAACCTTCCACGTCGGATCGCATC	3534
OY	1440	TGAAGTGTAAACGGTGTGTCTGTGAGCCAGCGCATATGCAATGTTCAAGTCAGGCGAT	1499
Db	3535	TCAAGTGTAAACGGTGTGTGTGTAAGCCAGCGTATAGCCATGTTCAAGTCAGGCGCAT	3594
OY	1500	TCAGAAGAAAGCAAAATTAAGATTAAGGCTTAAGGCGGGATATCTGAACCTTTATGA	1559
Db	3595	TGAGAAGAAAGCAAAATTAAGATTAAGGCTTAAGGCGGGATATCTGAACCTTTATGA	3654
OY	1560	TAAAGTGTGTGTCATCTGGACTTCGAAATCAATCAACTCAGGCAATTCCTGACATGCGA	1619
Db	3655	TAAAGTGTGTGTCATCTGGACTTCGCAATCAATCAACTCAGGCAATTCCTGACATGCGA	3714
OY	1620	TACCAATATTTTTTCAAGGCAACGATGTAACCAAGCGGTGATTAACGCCATATGACGATC	1679
Db	3715	TACCAATATTTTTTCTGTGACCAAGTGAACCAAGCGGTGATTAACGCCATATGACGCTC	3774
OY	1680	GTCAAAATCTTTTCCGTGATGACTGAACGAAATTCGTGTTAACCATACTCCGCTCTG	1739
Db	3775	GTCAAAATCTTTTCTGTGACTGAACGAAATTTTGTGTTAACCATACTCCGCTCTG	3834
OY	1740	TGTCCTCATGACCGGATCGCAAAAGCAGTGTATGCTCGACTGACTATGAGTGCAGAGATCC	1799
Db	3835	TGTCCTCATGACCGGATCGCAAAAGCAGTGTATGCTTGTGACTATGAGCCGCAAGATCC	3894
OY	1800	GAATGTTAAGGCTTAAGTCTCATCAGTTATATCATGGGAGACGATCCCAAGCTGTT	1859

Db		3995	GAAATGGTAAAGGCTGTGTCGCCCACTTTATACATGGAGAAGCATCTCCACAAGCTGTT	3954
Oy		1860	GGCGGTCCCCGACAAAAAAGAGCATTATATGTCTGTCGGGAGCGCAATTTTGAGATCTTT	1919
Db		3955	GGCGGTTCCCGACAAAAGAGGATTCGTGCTGCTCGGGAGCGCAATTTTGAGATCTTT	4014
Oy		1920	CCGGCGCTTTGCCAGACACTTATTTCTGCTGCGCTGCATTAAGAACCAAATATGTTATCCA	1979
Db		4015	CCCAGCGCTTTGCCAGATCTTAGTTCCTGCTGCGCTGCATTAAGAACCAAATATGTTATCCA	4074
Oy		1980	ACATGATTGGCTTACAGACAGGAATGCGGGGGAGCTTTAAACTCAACGGCGGTGTGA	2039
Db		4075	ACATGATTGGCTTACAGACAGGAATGCGGGGGAGCTTTAAACTCAACGGCGGTGTGA	4134
Oy		2040	GGATTTTTTATTTCTGAMAACTTTTCITTCAGACACTGAGCACGGCTAATGATACCGAGT	2099
Db		4135	GGATTTTTTATTTCTGAMAACTTTTCITTCAGAGCGCTGACATGCTAATGATACCGAGT	4194
Oy		2100	TTACTTGGCGGGTTGCAGTTGTTCTTCAACAGGTGATGSGTGGAGGGTGTCTATTACAC	2159
Db		4195	TTACTTGGCGGGTTGCAGTTGTTCTTCAACCGGGATGGATGGAGGGCGGTATTACAGC	4254
Oy		2160	CGCGTGAACGCGCTCTGTGCAATTAATCCAAATGAGAGGCACTTTGGCAAGGGCA	2219
Db		4255	CGCGTGAACGCGCTCTGTGCAATTAATCCAAATGAGAGGATTTTGGCAAGGGCA	4314
Oy		2220	TCCTCTGGAACACTCTTGTGGAAGAGATTAATACTACCGCAGTAGAAATTA	2267
Db		4315	TCCTCTGGAACACTCTTGTGGAAGAGATTAATACTACCGCAGTAGAAATTA	4362
RESULT 12				
LOCUS	ATU237588	29524 bp	DNA	linear BCT 09-JUN-1999
DEFINITION	Agrobacterium tumefaciens Ti plasmid pTiC58 T-DNA region.			
ACCESSION	AJ237588			
VERSION	AJ237588.1	GI:5042179		
KEYWORDS	acetoindoline synthase; agrocinopine synthase; IAH; isopenentenyl transferase; nopaline synthase; ornithine cyclodeaminase; protein 6a; protein 6b; torf1 gene; torf10 gene; torf11 gene; torf12 gene; torf13 gene; torf14 gene; torf15 gene; torf16 gene; torf17 gene; torf18 gene; torf19 gene; torf2 gene; torf4 gene; torf5 gene; torf6 gene; torf7 gene; torf8 gene; torf9 gene; tryptophan 2-monoxygenase.			
SOURCE ORGANISM	Agrobacterium tumefaciens Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.			
REFERENCE	1			
AUTHORS	Gielen,J., Terry,N., Van Montagu,M. and Villarroel,R.			
TITLE	Complete nucleotide sequence of the T-DNA region of the plant tumor inducing Agrobacterium tumefaciens Ti plasmid pTiC58			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 29524)			
AUTHORS	Gielen,J.			
TITLE	Direct Submission			
JOURNAL	Submitted (24-MAR-1999) Gielen J., Genetics, VIB/UG, KL			
FEATURES	Lederjansstraat 35 Gent, 9000, BELGIUM			
source	Location/Qualifiers			
	1..29524			
	/organism="Agrobacterium tumefaciens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:358"			
	/plasmid="pTiC58"			
	complement(1..575)			
gene	/gene="torf1"			
CDS	complement(<1..575)			
	/gene="torf1"			
	/function="unknown"			
	/codon_start=1			
	/transl_table=1			
	/product="hypothetical protein"			

```
gene
CDS
/protein_id="CAB44625.1"
/db_xref="GI:5042180"
/translation="MHPEDRARGLRRAFLRLHRCFLPSHPAQAAYHGDOLSPRVAINLF
ROEIRRPSRGQRTKIOEACLRMGCEBERBGLSEBERAFYAHADYICIFQOLIDQCR
OSITFTFEVDRLGRDHDPYAVRRKHGDTAGTARAIAMRADAPSSRRWYATPMTISG
RLVLVIGSATGSGMTNANSTGSSGAGRTS"
/length=604
/complement=604
/feature="torf2"
/gene="torf2"
/complement(604..1479)
/gene="torf2"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44626.1"
/db_xref="GI:5042181"
/translation="MVQSPSGICGRGAGATRPICEARFRTRNPTGRRRLQSGISASE
LEIIMNMIATGFNTASTDGLDELSELRLEAGVDTVELASSIDLJAGGRIVE
RGDLRLTLETFSPRYTVHGLVSNFMDPATRRHQVDAKALVEICSIDARVLVOHA
GFLRADQIARAGADEROCALLELEGEHAKAYGRIAPENFTTEPGQYOTPAEVAA
TVKAVNHPDVVALIDFKRLPGIQLSRPRPRSVARRGPIVLAGAPGRLSRTATRTIV
VTDGMECSGTWTS"
/length=805
/complement=805
/feature="torf3"
/gene="torf3"
/complement(805..1353)
/gene="torf3"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44627.1"
/db_xref="GI:5042182"
/translation="MIDGPRCGNFGSLTLARLRCENILKGDANVTRGMAQPEQ
RFAITLVSAQLSDILCAQAKARVLDENAGVDAAVNLQCCISDLPARCRVHEVRH
EAVHAGIEGRLLKKGSEPIATLFDPAATCQIEAGKEFPRVDADCKAQLSEAVK
LAVRRCRIEARCNIIVNDFOF"
/length=1359
/complement(1359..1886)
/gene="torf4"
/gene="torf4"
/complement(1359..1886)
/gene="torf4"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44628.1"
/db_xref="GI:5042183"
/translation="MPGEPAQOTALARLANVADKLMTGEIRTOYALCYRRTDSSGP
LEVLITSRGTGWVLPKGMMAKKKPHVEVASQAEAGVGRGVRKKAHGYTVYKR
LGGQEFIPAMVOVHLIDVQRMEDDFPRHQREIQWSPRLAASAVGEPRLRGLFARLE
ERLQRAAVASKAG"
/length=2090
/complement(2090..2644)
/gene="torf5"
/gene="torf5"
/complement(2090..2644)
/gene="torf5"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44629.1"
/db_xref="GI:5042184"
/translation="MGESGEIPGASGRVWVVIDPTFNFVRGQNMALISGLYENR
RPSFGVIYAVRDILVGKTPVPLNGOPIKPLPLMDSRASTGSRPLSDLKISF
RFGGAATLSMIEVMGERTGVSLGDSMTVMMAALPLIGHLISNTIDMDRIDLSKTL
RFGAGSEERLSKRPILIEKVSAAA"
/length=3623
/complement(3623..3647)
/rpt_family="T-DNA left border"
/rpt_type=DIRECT
/complement(4009..5124)
/gene="torf6"
/gene="torf6"
/complement(4009..5124)
/codon_start=1
/transl_table=11
/product="agrociniopine synthase"

gene
CDS
/protein_id="CAB44630.1"
/db_xref="GI:5042185"
/translation="WYNICKGVOECETSTSLTLGCMERNINSELDITLTSNDLPVVS
HDTNTRISTLPRLRLIELHSADVKNIPVIRKVSNGELTESITVNDIIPLEPLD
KQVDPVAGETAFLLDGRDFEAAHITLNLSSRPKVRQVVLLEFYFEYSSGAEPDAVQ
VSEBPMRETAVALMPVTFPELPRKLKLDLSDIAVEDYEGEKIMDSMLRQPMRYE
AVHITVARVKEPEQLGCDKPEIVRANADYAAVRLAYVYKDNQNEVVRMRPHLKATAT
RCYDFSAOLENGEKAFFSINIRTRGPMPEPTEDEKRYIRGCGYIGNAALAIADVISDR
SEDEMSFWEMRNKVRRRVDHCHCPHLDVLEQDQKSP"
/length=6011
/complement(6011..6694)
/gene="torf7"
/gene="torf7"
/complement(6011..6694)
/gene="torf7"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44631.1"
/db_xref="GI:5042186"
/translation="MYERPPVFDIVDTWEIREREVLKPVLSVAEDYCYFVONTVASA
OREWDLVNLNLPSPDSIDDLNLRPADAPCLHGPLLEPNLILKSPMYVYCSPEEM
RSCATRRFTEGISNRGVICTVPPYABGTRDDMSWQNOACVNTCSGRKDLAYIAF
LFTSSLQESSYWHTKNGIYSFLAPSOTDAFCCEILCVGRALFEDRSRKEKLRNCLKL
LNRYRLKLP"
/length=7150
/complement(7150..8721)
/gene="torf8"
/gene="torf8"
/complement(7150..8721)
/gene="torf8"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44632.1"
/db_xref="GI:5042187"
/translation="MENQTPVEALIERLKSSWLSLHVRKILPSPFLGSPYSYIE
TVQGDQADICRHSVTPLPEDDEGMSVSHHQIIFCSYHOPADHVDPAFLRAAPBAY
IPDEDIADYSVPFASSENSLTDILMRMPLSTRNANATQARKERDDAFVYLEPSP
KEYKYLRRSAOSIOFOEMRMOKCHLPLMBEGLFWMLVIMQIAVLDVIDVASIGA
EVAVALRSGPITLOTCLMVLAIDYPRADTRLAAHLACBVLGSLPVILIGLFRILAMPQ
HEGTTFRHAKDVAMKASLCLTNOMLQNRKNIRDLKYSFLRKRIISTQSSDFTSEY
TRLRNSSTIIYSMLDTFEVSNRLRIQIGHKQDHEDALVLRWVGHIADONDSPSGG
KAGILVAIITMAGLICLSTFPDPAVSGLLAANLVPLAARAADVMDNSQSHDLRL
FLTAGVSPFSTLFLVJGNSSWMLLTKKSFLAGRANIHFWIAFVLIFPLSLPRLMGRL
FMVYSGRLNDRSRA"
/length=9762
/complement(9762..10493)
/gene="torf9"
/gene="torf9"
/complement(9762..10493)
/gene="torf9"
/function="phosphatase"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44633.1"
/db_xref="GI:5042188"
/translation="WHTSLMSYNONMIPSGDPRVPVPTIDLFPANTSLEELQOGLCAAV
ANYRDLTRTKILPFOOSWYVRVLSGGLDEETLRATKSGVFPCCSILPGANODRIMO
HDSPIYITSLNLSQYVEGKTLRLCYRFRVYLAASSIPRTTIPAEELKNVANSILPG
AISMPIAVSCFVAIILPSKKFTISMSKITTRGAEHEHFFVFSREENDIIPYDIVAVKXT
LFSQSDSEFGLPNLSFDELONLIST"
/length=11195
/complement(11195..11890)
/gene="torf10"
/gene="torf10"
/complement(11195..11890)
/gene="torf10"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44634.1"
/db_xref="GI:5042189"
/translation="MNVYQRPVTVISLIRINREBAKVLIGAVVYVRNFEVQTLAD
AQKNWVKSLLVLDPDGDATVGLITFWSRYACLHGRPLGLDTIAVNDPLVLYICPRRK
LEEYPERIVSFHSEIGSVCSMPEDAGVTEKRYVGNLISPGSCLPDLALEAVVA
FLPSKSLKLPYSVYEVVHNDRYVHKFPALLPSSRPHFEVVAAGLAVPAKKKPSGCI
LACRTGKNTCL"
```

gene 12368..13312
/gene="torf11"
12368..13312
/gene="torf11"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB44635.1"
/db_xref="GI:5042190"
/translation="MNVPRPGRAPISLESINDMEMHFLTSACILHYIYIDNLLI
YOEIWKYWRGLAESALKNELIALYLRPPRPTLQSIABEVLCKDAPIYVCS
REMLSHVOSKDLCOFGRANLACTLAPLYRGTIEVNRCLFRLRGNREVTGSE
DLSEFVAFLEPTTAFRRPRPYLGENVQFLREGLELPGHILAFGLRSEILADGS
SMWKSIVGESSRLNHVAVSVELSRNVPTAETDNNMAALGTDVAVPRPRLTPQD

Query Match 88.9%; Score 2016.2; DB 1; Length 29524;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 2137; Conservative 0; Mismatches 128; Indels 4; Gaps 4;

QY 1 ATGTCAGCTTCACCTCTCTTGATACAGCGGATATCTCCCAACCAAAATGGTGGAT 60
Db 19987 ATGTCAGCTTCAGCTCTCTTGATACAGCGGATATCTCTACCAAAATGGTGGAT 20046
QY 61 CTGACATGGTTCGATTAAGCGCGATGAATTGGACCGAGGGTTTCCAGTGCCTTTAGAA 120
Db 20047 CTGATTAATGGTTCGATTAAGCGCGATGAATTGGACCGAGGGTTTCCAGTGCCTTTAGAA 20106
QY 121 CGAAGACCTTCTAGGGGAGAGAGATTACTCAAACTTCCACCGAGTGCACGCTGGGTTA 180
Db 20107 CGTGAACCTTCTAGGGGAGAGAGATTACTCAAACTTCCACCGAGTGCACGCTGGGTTA 20166
QY 181 GCTTGAAGAGCGCGCGATGGTCTGCCCGAGATCTCAGCTGGTGAAGATGCA 240
Db 20167 GCTTGAAGAGCGCGCGATGGTCTGCCCGAGATCTCAGCTGGTGAAGATGCA 20226
QY 241 GTTCTCTCCGCTTATATCTAATTGGCAAGAAATTCGGGCGCGATCTTGAATCGAA 300
Db 20227 GCCTCTCCGCTTATATCTAATTGGCAAGAAATTCGGGCGCGATCTTGAATCGAA 20286
QY 301 CTTTGGGCGCGGACAGATGATGCTCTGTTGCCATTCGACTTGAACACCATTTGCATG 360
Db 20287 CTTTGGGCGCGGACAGATGATGCTCTGTTGCCATTCGACTTGAACACCATTTGCATG 20346
QY 361 GATTTCCGAGACAACTAATCCAGCCCTGTTTTGCTGAGCGGTAAGATGCA 420
Db 20347 GATTTCCGAGACAACTAATCCAGCCCTGTTTTGCTGAGCGGTAAGATGCA 20406
QY 421 CCGATTGATCTAATCTAATCTGAGCAATCTCTAAGACTGCCGCTTTGCAACC 480
Db 20407 TCCAGCGATCTTATGATCTTCTGAGCAATCTCTAAGACTGCCGCTTTGCAACC 20466
QY 481 CTGCGAATGCCGCTGTAAGAGATGCAAGTAATGCGTTACCGGGTTTACCATAAC 540
Db 20467 CTGCGAATGCCGCTGTAAGAGATGCAAGTAATGCGTTACCGGGTTTACCATAAC 20526
QY 541 CTTGAAGGGGCGCGCAATTTGACATGATGCTTAATGTTGAACCTGATGCAAGGCT 600
Db 20527 CTTGAAGGGGCGCGCAATTTGACATGATGCTTAATGTTGAACCTGATGCAAGGCT 20586
QY 601 TCGGCAAGTTCCTTTCCAAACATGCACTTGTCTTACAGACTGACAGCCGTTTTTGACCA 660
Db 20587 TCGGCAAGTTCCTTTCCAAACATGCACTTGTCTTACAGACTGACAGCCGTTTTTGACCA 20646
QY 661 TGTTCGATAGTGAAGCGATCGGCTTTTCCGAGGATGTTCTTAAGCGAAAGTGGCG 720
Db 20647 TGTTCGATAGTGAAGCGATCGGCTTTTCCGAGGATGTTCTTAAGCGAAAGTGGCG 20706
QY 721 GTCAATGGCGCTGCAATTTCCGACATGCTGTGTGGCAACGAACTGCTTACGCTGGGTA 780
Db 20707 ATCAATGGCGCTGCAATTTCCGACATGCTGTGTGGCAACGAACTGCTTACGCTGGGTA 20766
QY 781 GACGATGTACAATATATGAAGCAATGATCGTGTGGAGCAAGCTTTGGTCAATGCT 840

Db 20767 GACGATGTACAATATATGAAGCAATGATCGGTTGGAGGCAAGCTTTGGTCAATGCT 20826
QY 841 TTCAGGAGCGCTCCTATGTCGCGGCGCAAAATGGGGGCGCATGGCATTTCTCTGCTGCA 900
Db 20827 TTCAGGAGCGCTCCTATGTCGCGGCGCAAAATGGGGGCGCATGGCATTTCTCTGCTGCA 20886
QY 901 TTCGCTGTTGTTTCTTCTCTGAGCGCTTACGAGCTCTCTTCAATGAGCCGTTCCAAAT 960
Db 20887 TCCTGCTGTTGTTTCTTCTCTGAGCGCTTACGAGCTCTCTTCAATGAGCCGTTCCAAAT 20946
QY 961 CCGGCAAGTGGACACTTACTTGTCTCAACAGGCGCTCCAAATCATGTGAAAGCCGGG 1020
Db 20947 CCGGCAAGTGGACACTTACTTGTCTCAACAGGCGCTCCAAATCATGTGAAAGCCGGG 21006
QY 1021 CAGCTGCCACCGAAGCTTTCATTCGCGTTTACAAGGTTGGGCGCTTTGAAAGGAC 1080
Db 21007 CAGCTGCCACCGAAGCTTTCATTCGCGTTTACAAGGTTGGGCGCTTTGAAAGGAC 21066
QY 1081 GGTTCATGAGCGGAGATATGTTGGCTTCGCTGCTGCTTATCTCAGGCTTGAA 1140
Db 21067 GGTTCATGAGCGGAGATATGTTGGCTTCGCTGCTGCTTATCTCAGGCTTGAA 21126
QY 1141 TCAGACATTAAGTGGGCTCATGACTCTGGCAAAATTTGGCTGAAACGTTTGGGAGG 1200
Db 21127 TCAGACATTAAGTGGGCTCATGACTCTGGCAAAATTTGGCTGAAACGTTTGGGAGG 21186
QY 1201 GATCTCTTCTCTCAGGATTAAGAGAGATCTTCTGGGCAACATCTCTGCTGGT -AA 1259
Db 21187 GATCTCTTCTCTCAGGATTAAGAGAGATCTTCTGGGCAACATCTCTGCTGGT 21246
QY 1260 ACATGAGATTTTCTCATGATTTGGGACTTATCAAGCTAATGGGAATAGGATCTGGCGG 1319
Db 21247 ACATGAGATTTTCTCATGATTTGGGACTTATCAAGCTAATGGGAATAGGATCTGGCG -GG 21305
QY 1320 GTTGTGTCAGTTTGAAGCGGGTTTATGAGATCTCTCGTGTCTCAACGAGATA 1379
Db 21306 GTTGTGTCAGTTTGAAGCGGGTTTATGAGATCTCTCGTGTCTCAACGAGATA 21365
QY 1380 TGAAGAAATACAGCGGATGCTCCCTGAAGAAATCTGAAGAACTTCCAGTCCGATC 1439
Db 21366 TGAAGAAATACAGCGGATGCTCCCTGAAGAAATCTGAAGAACTTCCAGTCCGATC 21425
QY 1440 TGAAGTGTAAACGATGCTGTGTGAGCAGCGCATATGCCATGTTCAAGTCAAGGCGAT 1499
Db 21426 TGAAGTGTAAACGATGCTGTGTGAGCAGCGCATATGCCATGTTCAAGTCAAGGCGAT 21485
QY 1500 TCAGAGGAAAGACAAATAAAGATTAAAGCTTAAGACCGGATATCTGAATTTATGA 1559
Db 21486 TCAGAGGAAAGACAAATAAAGATTAAAGCTTAAGACCGGATATCTGAATTTATGA 21545
QY 1560 TAAAGTGTGTCACTCTGAGACTCGCAATATCCAACTCAGGCAATTCGCTGACATGCGA 1619
Db 21546 TAAAGTGTGTCACTCTGAGACTCGCAATATCCAACTCAGGCAATTCGCTGACATGCGA 21605
QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACGACCATATGACAGATC 1679
Db 21606 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACGACCATATGACAGATC 21665
QY 1680 GTCAAAATCTTCTCTGATGATGAGCAAAATTTGTTAGACCATATCTCCGCTTG 1739
Db 21666 GTCAAAATCTTCTCTGATGATGAGCAAAATTTGTTAGACCATATCTCCGCTTG 21725
QY 1740 TGTCTCTATGAGAGGATGGCAAAACAGATGATTCCTCTGACATTAAGTCCGAGATCC 1799
Db 21726 TGTCTCTATGAGAGGATGGCAAAACAGATGATTCCTCTGACATTAAGTCCGAGATCC 21785
QY 1800 GAATGTAAAGGCTATGATGCTCATGATTAATACATGAGAGAGCATCCCAACAGCTGT 1859
Db 21786 GAATGTAAAGGCTATGATGCTCATGATTAATACATGAGAGAGCATCCCAACAGCTGT 21845
QY 1860 GCGGCTCCCGACAAAGGATTAATGCTGCTGCGGAGCAATTTGAGATCTTT 1919

Query Match 82.9% Score 1880.8 DB 1: Length 10200;
Best Local Similarity 89.9% Pred. No. 0;
Matches 2039; Conservative 0; Mismatches 227; Indels 2; Gaps 2;
QY 1 ATGTGACCTTCACTCTCTTGTATTAACAGTGGCATCTCCCAACCAAAATGTGTAT 60
DB 4657 ATGTGACCTCACTCTCTTGTATTAACAAATGGCATCTCTACCAAAATGTGTAT 4716
QY 61 CTGACATGTGTGATTAAGGCGGATGATTTGACCCGAGGGTTTCCATGCTTCTTAGAA 120
DB 4717 CTGATATATGTGTGACAAAGGCTGATGATGACCGCGGCTTCCGAGCTTCTAGAA 4776
QY 121 CGAAGACCTTCTAGGGGAGAGAGATTAATCAAGCTCCACCGAGTSCAGCGCTGAGT 180
DB 4777 CGAGAACCTTCTAGGGGAGAGAGATTAATCAAGCTTCCGCGAGTSCAGCGCTGAGT 4836
QY 181 GCTTGCAGAAAGGCTGCGGATGATGCTTCCGAGATCTCACTGTGTGAAAGTAGCA 240
DB 4837 GCTTGCAGAAAGGCTGCGGATGATGCTTCCGAGATCTCACTGTGTGAAAGTAGCA 4896
QY 241 GTTCTCTCCGCTTATATCTATTTGAGCAAGAAATTTGAGGCGGATTAATGATGCA 300
DB 4897 GTCTCTCTCCGCTTATATCTATTTGAGCAAGAAATTTGAGGCGGATTAATGATGCA 4956
QY 301 CCTTGGGCGGCGGCAAGTGAAGTGTCTGCTTCCATCGACTTGCACACATTTGCATG 360
DB 4957 GCTTGGGCGGCGGCAAGTGAAGTGTCTGCTTCCATCGACTTGCACACATTTGCATG 5016
QY 361 GATTTCTCCGAGACCACTAATTCAGACCCCTGTTTTGCTGAGCGGTAAAGATGTGA 420
DB 5017 GATATTTCCAGAGCTCAACTTCTCCCAACCATGAATTTGCTGAGTGTAAAGATGTGA 5076
QY 421 CCGATTGATCTTATGCTATTTCTGTGGCAATTTCAATCTCTAAGACTGCGGCTTTCCAGC 480
DB 5077 CCGAGCACTTATGCTATTTCTGTGGCAATTTCAATCTCTAAGACTGCGGCTTTCCAGC 5136
QY 481 CTGCGAATGCGGCTGTGAGAAATGGCAGATGAATGGCTTACCGGGTTTACCATACC 540
DB 5137 CTGCGAATGCGGCTGTGAGAAATGGCTTACCGGGTTTACCATACC 5196
QY 541 CTGGAAGGCGCGTGCATTTGACATGTAGCTTATGTGCGAAACCTGATGCTGAAGGT 600
DB 5197 ATTGAAGGCGCGTGCATTTGACATGTAGCTTATGTGCGAAACCTGATGCTGAAGGT 5256
QY 601 TCGGCAAGTTCCTTTCCAACTGATCTGTGCTGACATCTGACAGACCGTTTTTGAACAA 660
DB 5257 TCGGCAAGTTCCTTTCCAACTGATCTGTGCTGACATCTGACAGACCGTTTTTGAACAA 5316
QY 661 TGTTCGAGTGTGAGCGGATCGGCTTCTTCCGAGGATGTTCTTAAGCGGAAAGTGGC 720
DB 5317 TGTTCGAGTGTGAGCGGATCGGCTTCTTCCGAGGATGTTCTTAAGCGGAAAGTGGC 5376
QY 721 GTCAATGCGCTGTGATTTCCGACTCGTGTGCGAAACGAACTGCTTCAATGCTGGGTA 780
DB 5377 GTCAATGCGCTGTGATTTCCGACTCGTGTGCGAAACGAACTGCTTCAATGCTGGGTA 5436
QY 781 GAGCATGTACATATATGAAGCAAGTGTGTGTGAGGCAAGCTTTGTGTCAATGCT 840
DB 5437 GAGCATGTACATATATGAAGCAAGTGTGTGTGAGGCAAGCTTTGTGTCAATGCT 5496
QY 841 TTCAAGGAGCGCTCTAGTGTGTGCGGAAATGGGGGCGATGCGATTTCCCTCGCTGCA 900
DB 5497 TTCAAGGAGCGCTCTAGTGTGTGCGGAAATGGGGGCGATGCGATTTCCCTCGCTGCA 5556
QY 901 TTTGCTGTTTTTTCTTCTCTGAGCGTTACGCGCTGTCTTCTGATGAGCGGTTTCCCAAT 960
DB 5557 TTTGCTGTTTTTTCTTCTCTGAGCGTTACGCGCTGTCTTCTGATGAGCGGTTTCCCAAT 5616
QY 961 CCGGCAAGTGCACCTTACTTGTGTCAACAGCGCTCAATACATGTGAAAGCCGG 1020
DB 5617 CCGGCAAGTGCACCTTACTTGTGTCAACAGCGGTTCGATCATGTGAAAGCCGG 5676

QY 1021 CAGCTGCCACCGAAGCTTTCCATCGCTTTACAAAGTGTGGCGCTGTTTGAAGAC 1080
DB 5677 CAGACGCCACCGAAGTGTTCATCGCTTTACAGGGGGGCGATCGTTCTTTGAAGAC 5736
QY 1081 GGTTTTCATGAGGAGATATTTGTGTGCTTGGCTGTGCTTATTACTAGGCTTGA 1140
DB 5737 GGTTCCTTGAAGGAGATATTTGTGTGCTTGGCTGTGCTTATTACTAGGCTTGA 5796
QY 1141 TCAGACACATTAAGGCGCTCATGCTCTGCGAAATTTGGCTGAACGTTTGGGAG 1200
DB 5797 TCAGGAGACATTAAGGCGCTCATGCTCTGCGAAATTTGGCTGAACGTTTGGGAG 5856
QY 1201 GAGTCTTCTCTTCAAGGATAGAGAGATCTTTGAGGACACATCTCTGAGG- TAA 1259
DB 5857 GAGTCTTCTCTTCAAGGATAGAGAGATCTTTGAGGACACATCTCTGAGGAGAA 5916
QY 1260 ACATGAGTTTTCTCATGATTTGGGACCTTATTCAGCTAATGGGAATAGATTTGGCGG 1319
DB 5917 ACATGAGTTTTCTCATGATTTGGGACCTTATTCAGCTAATGGGAATAGATTTGGC- GG 5975
QY 1320 GTTGTGTCAGTTTTTGAAGCGGTTTATTTAGATCTCCGCTTGTCTCAACGATTA 1379
DB 5976 GTTGTGTCAGTTTTTGAAGCGGTTTATTTAGATCTCTTGTGTCTAATAAGATA 6035
QY 1380 TGAAGAAATCAAGCGATGTGCTTGAAGAAATCTGAACTTCAAGTGCATC 1439
DB 6036 TGAAGAAATCAAGCGATGTGCTTGAAGAAATCTGAACTTCAAGTGCATC 6095
QY 1440 TGAAGTGTTAACGGTGTCTGTGAGCCAGCATATGCTCATGTTCAAGTGCAGGCGAT 1499
DB 6096 TCAAGTGTTAACGGTGTCTGTGAGCCAGCATATGCTCATGTTCAAGTGCAGGCGAT 6155
QY 1500 TCAGAGGAAAGACAAATATTAAGATTAAGGCTTGAAGCGGATTAATCTTATGA 1559
DB 6156 CGAGAGGAAAGACAAATATTAAGATTAAGGCTTGAAGCGGATTAATCTTATGA 6215
QY 1560 TAAAGTGTGTGCTCATCTGAGCTCGCAATATCAACTCAGGACTTGCCTGATGCA 1619
DB 6216 TAAAGTGTGTGCTCATCTGAGCTCGCAATATCAACTCAGGACTTGCCTGATGCA 6275
QY 1620 TACCAATATTTTCAAGCACAGTGAACCAAGCGTTGATTAAGCCATATGAAGATC 1679
DB 6276 TACCAATATTTTCAAGCACAGTGAACCAAGCGTTGATTAAGCCATATGAAGATC 6335
QY 1680 GTCAAACTCTCTGATGCTGAAGAAATTTCTGTTAGACCATATCTCCGCTTGG 1739
DB 6336 GTCAAACTCTCTGATGCTGAAGAAATTTCTGTTAGACCATATCTCCGCTTGG 6395
QY 1740 TGTCTCATGAGCGGATCGCAAAAGCAGTATTCCTGAGCTATGATGCGAGATCC 1799
DB 6396 TGTCTCATGAGCGGATCGCAAAAGCAGTATTCCTGAGCTATGATGCGAGATCC 6455
QY 1800 GAATGTAAAGTGTAGTGTCTCATGCTTATACATGAGAGAGACTCCCAAGCTGTT 1859
DB 6456 GAATGTAAAGTGTAGTGTCTCATGCTTATACATGAGAGAGACTCCCAAGCTGTT 6515
QY 1860 GGGGTCCTCCGCAAAAAGAGGATTAATGTCTGCTGCGGAGCGCAATTTGAGATCTT 1919
DB 6516 GGGGTCCTCCGCAAAAAGAGGATTAATGTCTGCTGCGGAGCGCAATTTGAGATCTT 6575
QY 1920 CCGGCGTTTTCCAGACCTATTTCTGCTGCGTGTATTAAGACCAAAATGTTATCA 1979
DB 6576 CCGGCGTTTTCCAGACCTATTTCTGCTGCGTGTATTAAGACCAAAATGTTATCA 6635
QY 1980 ACATGATGCTTACAGAGAGATGCGGGGAGCTTTCAAACTCAACGCGGTGTGA 2039
DB 6636 ACATGATGCTTACAGAGAGATGCGGGGAGCTTTCAAACTCAACGCGGTGTGA 6695
QY 2040 GGAATTTTATTTGAGAGAACTTTTCTTCAAGACTGGAACAGGCTTAATGATCCGAGT 2099
DB 6696 GGAATTTTATTTGAGAGAACTTTTCTTCAAGGCTGGAACAGCTTAATGATCCGAGT 6755
QY 2100 TTAATTTGCGGTTGAGATTTCTTCAAGGTGATGGGTGAGGCTTATTCAGAC 2159

Db 6756 TTACTTGGCGGGTTCAGATTGTCTCTTACACCGTGGATGGGTGAGGCGCTATTTCAGAC 6815
Oy 2160 CGCGTGAAGCCGCTCTGTGAATTATCCAAATTTGCGAGGCAATTTTGGCAAGGGCAA 2219
Db 6816 CGCGTGAAGCCGCTCTGTGCGCAATTATCCAAATTTGGAGGCAATTTTGAAGGAGCAA 6875
Oy 2220 TCCTCTGGAACACTCTTGAAGATATTAATACCGAGTGAATTA 2267
Db 6876 TCCTCTTAACACCCCTTGGAGAGATATACTACCGCAATTAATTA 6923
RESULT 14
LOCUS AVU83987 15463 bp DNA linear BCT 02-MAR-2001
DEFINITION Agrobacterium vitis plasmid pT1m4 agrocinopine synthase and (5)
genes, (1aah) pseudogene, and (1aah), (1pt), (6b) and (ocs) genes,
complete cds.
ACCESSION U83987
VERSION U83987.1 GI:1814322
KEYWORDS Agrobacterium vitis (Rhizobium vitis)
SOURCE Agrobacterium vitis
ORGANISM Agrobacterium vitis
REFERENCE 1 (bases 1 to 15463)
AUTHORS Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
TITLE Sequence and functional analysis of the shoot-inducing octopine
T-DNA of Agrobacterium vitis strain CG474
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 15463)
Osten, L. and De Ruffray, P.
REFERENCE 2 (bases 1 to 15463)
Osten, L. and De Ruffray, P.
AUTHORS Direct Submission
TITLE Submitted (07-JUN-1997) Institute for Plant Molecular Biology, rue
JOURNAL du General Zimmer 12, Strasbourg 67084, France
FEATURES
source
1. 15463
/organism="Agrobacterium vitis"
/mol_type="genomic DNA"
/strain="Tm4"
/db_xref="taxon:373"
/plasmid="pT1m4"
3857. 6588
/organism="Agrobacterium vitis"
/mol_type="genomic DNA"
/db_xref="taxon:373"
/insertion_seq="IS866"
15438. 215463
/organism="Agrobacterium vitis"
/mol_type="genomic DNA"
/db_xref="taxon:373"
/insertion_seq="IS867"
139. 15310
/note="shoot-inducing octopine T-DNA"
139. 165
/note="flanking repeat; left border of the T-DNA"
complement (373. 1629)
/codon_start=1
/transl_table=11
/product="agrocinopine synthase"
/protein_id="AAB41872.1"
/db_xref="GI:1814323"
/translation="MEGTAMPATLTHMDLPENTTPPEVLARHSVPLKLERLENTLPLQ
VIAHGMFNGNRIQECTASSLLAALGQGRNISELDVCLTSNIPVSHDLTWVS
EKXGDKFENIHSSDINNVPVIRVSNGLIQDYREITDHPILDEILGKVSAPD
ATIFLDGRNEAIVAVLSHRPVYQRAVLFYTFEPDGAFAVDVILTAQPSAPMR
KSIAPALPPEELCLARLRQVSEPTVDVLAKAMIDSLMODMRVAVAVVPSF
VSNRLDRAVVDKVLAFSDQAAVRLAHYKEDIMTKARPLKFAATVTRCQFPAAL
LNGERAERSIDINTGRARPHETDERKHIMRKGTGNSKATIDAVTISDRSEDMATW
EMNQGIDREVSLSPLDVDVGTSE".
gene
2235. 2918
/gene="5"
CDS
2235. 2918

/gene="5"
/codon_start=1
/transl_table=11
/protein_id="AAB41873.1"
/db_xref="GI:1814324"
/translation="MDHSRPIFNVIDCSHIOQRRELEIVLLATRRAYLRFAPQDLPA
QRSMIDSIIDTNVIDPAIDEVYKFRSVACLPAPVGMISLINTSMLVYCSFDM
RRACRPFYEGSGKGVISTVPYVQGITKEMRCOMENVECHITNNEKNDLPAYLAF
LPRTSIQNTSFSHVXIGGDAFLAPSRADPFCEIVAVAGRALFODNKLKAPKAGWSMA
LISLRLV"
complement (3403. 7530)
/gene="1aah"
/gene="1aah"
complement (3403. 3856, 6589. 7530)
/note="1aah"
/note="interrupted by the bacterial insertion element
IS866"
/pseudo
/codon_start=1
/transl_table=11
3857. 3864
/note="flanking repeat; left border of the IS866 insertion
element"
/rpt_type=direct
6581. 6588
/note="flanking repeat; right border of the IS866
insertion element"
/rpt_type=direct
7868. 10135
/gene="1aah"
7868. 10135
/gene="1aah"
/codon_start=1
/transl_table=11
/protein_id="AAB41874.1"
/db_xref="GI:1814325"
/translation="MSASSLIDKQDHFSTKIVLIMVDAKDELRVAAFSERES
RERRISQIGECNAGLACRLADRPESISAGRVAVLSAYIVGSEILRWILEPERS
VTRVRSYVAIILAPSCMDISRAQLQTNLUSGRCAESDLSHPTAISSETARSRT
LQAPYBESLKSVTGFTVITIEAVPFDWVAGRNMLKASQSFPTIDLVYRLEPL
DKCSDSGRIGFPEDEVPRKAVITGIGLVVASELSLRAGVDVTIYEGDVGGL
WSHAFDAGVAEMGMRPPAASCLFFLEERYGLSNRPFPNPGTVDTLVYEGCR
VMKAGQOAPPLFHYRYSGMHAFKQGFLEGLVYGLSPAITALKSGDIRRAHDSMO
IWLNRGRESFSAIRIFLGTHPGGETSWPDMDFKMGISGGGPGVPESGFT
EILRLVINYENQRCSEGISLPRLASOVYNGVSQRIRHQVYRIEKRKTIK
IRKSGISELYDKVVTSGLANIQKHCLCTDTTIRAVNQAVDSHHTGSKLPL
TERKWFDMPLSPCVLMDGFAKAVCYCLADYEDDPNGKGLVLSYTNEDSHKLLAVPD
KXERLCILRLDAISKSPFPAOHLPACADYDQNVQHDWLTENAGSRFKLNRGDF
VSELEFQALDITNDTGYVLAGSGSCFTGAWGEGAIQTACNAVCATIHNCGGLAAND
PLKHPKRYNVRNRN"
10824. 11546
/gene="1pt"
10824. 11546
/gene="1pt"
10824. 11546
/gene="1pt"
/codon_start=1
/transl_table=11
/protein_id="AAB41875.1"
/db_xref="GI:1814326"
/translation="MDLRLIFGPTCTGKTSTAIALAQGTGLPYLSIDRVQCCPOLSTG
SGRPVYEBLGTTRLYLDDRLPVKGIIAEQAHRLIABVYVNEAGGLIEGSGISL
LCKMQSGWASADFRRHIIIRHKLADETEMKAARVKOMLCPAIGPSLIOELVLYMN
EPLRLPIITKEIDGYVYAMFASONRTTPMLQLDADMGKLIHGAQGYLTHAROE
HEPPVSAAPFGFEGPFGAY"
complement (12224. 12847)
/gene="6b"
complement (12224. 12847)
/gene="6b"
/codon_start=1
/transl_table=11
/protein_id="AAB41876.1"
/db_xref="GI:1814327"
/translation="MTVANMQRVDFTRLINAGELQRLQEARTDGALLAEIVYFHP
GATPEGDDEYILTGGLVYVYVSEGTALNRLPLPNSNSNGFVTVAIPLMLDGT
QTLNLTLORCQDGGI VNVYHSGRTNEPFLA LMLNSNCFVRFGTDELNGASYGFIAR

Db	9667	GAACGGTAAAGTGTGGTGGCTCATCACTTATATATGGAGAGAGACTCCACAGACTATT	9726
QY	1860	GGCGGTCCCCGACAAAAGAGCGATTATGTCTGTCCGGAGCGAATTTGAGATCTTT	1919
Db	9727	GGCGGTCCCCGACAAAAGAGAGATATATGTCTGTCCGGAGCGAATTTCAAAATCTTT	9786
QY	1920	CCCGGGCGTTTGGCCAGACCTATATTTCTGCTGTGGCTGATTACAGACCAAAATGTTATTTCA	1979
Db	9787	CCCGGTGTTTTGGCCAGCATCTAGTTCCTGGCTGTGGATTACAGCAAAATGTTGTTCAC	9846
QY	1980	ACATGATTGGCTTACACAGAGATGCGGGGGAGCTTCAAACTCAACCGGTGTGTACA	2039
Db	9847	ACATGATTGGCTTACACAGCAGAGATGCGGGGAGCGTTTCAAACTTAACCGAGCGGGCGCA	9906
QY	2040	GGATTTTATTTCTGAGAACTTTTCTTTCAAGACACTGAGACACGGCTATATGATACCGGAGT	2099
Db	9907	AGATTTTATTTCTGAGAACTTTTCTTTCAAGCGCTGAGACAGACTATATGATCTGAGAT	9966
QY	2100	TTACTTTGGCGGGTTGACAGTTGTTCCTTACACAGGTGAGTGGGTGAGGGTGTCTATTCAGAC	2159
Db	9967	TTACTTTGGCGGGTTGACAGTTGTTCCTTACACCGGTGAGTGGGTGAGGGCGCTATTCAGAC	10026
QY	2160	CGCGTGTAAACGGCGTCTGTGCAATTTATCCAGAAATGTGTGAGGCAATTTTGGCAAAAGGGCAA	2219
Db	10027	CGCGTGTAAACGGCGTCTGTGCAATTTATCCAGAAATGTGTGAGGCAATTTTGGCGAAGAGCAA	10086
QY	2220	TTCCCTCGAACACTCTTGGAGAGATATTAACCTACCGCAGATAGAGAAATTA	2267
Db	10087	TTCCCTTTAAACAACCTTGTGAGAGATATTAACCTACCGCAATATAGAGAAATTA	10134

```

RESULT 15
AVU83986
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 14960
/organism="Agrobacterium vitis"
/mol_type="genomic DNA"
/strain="CG474"
/db_xref="taxon:373"
/plasmid="PT:CG474"
7244. 8395
/organism="Agrobacterium vitis"
/mol_type="genomic DNA"
/db_xref="taxon:373"
/insertion_seq="IS870-like insertion element"
11921. 12767
/organism="Agrobacterium vitis"
/mol_type="genomic DNA"
/db_xref="taxon:373"
/misc_feature
252. 14745

```

repeat_region	note="shoot-inducing octopline T-DNA"
252..278	/note="Flanking repeat, left border of the T-DNA"
/tpc_type=direct	
complement (496..1854)	
/codon_start=1	
/transl_table=11	
/product="agrocinnopline synthase"	
/protein_id="AAB41866.1"	
/db_xref="GI:1814316"	
/translation="MALNVRVLYTLLTMTCHNIPPEFAOVLQCRPKHKEGVTPATPTLHMDELPEVTFDLVAHSVRPLRELEBDPLQVTEIRGMPFNIGKIOGCTASLLA	
GLGQGRRLSMDVCLTSDNVAIYSHDLNTWRVSEKGDCLKFNEIHSKIKDVTYIR	
EVSNTIYDCKTFETVDTHPSLTELGVKFLANPDAITFLGGRNTEAHVIAWLSHQET	
YHORVNIYFYFEYFHGCAFDVAALKAQPSADMRKSIALMPALFPEELCELAIRPOT	
EPYVDLILACKAWIDSMLOQMDRIVAHVVFSGVTNNLLGOVVDKVLAFQSDAEA	
VLRYLYVEDMELNIRAKRPHLKEFAVTVTCYDFPALDCCGESEFIDIKTGARSHED	
ERKIRKRMKGRPGNSATIDAVTISDRPEDEMAITEMRNQIGIDREVSILSHLDVSVET	
PG"	
2363..3046	
/gene="5"	
2363..3046	
/gene="5"	
/codon_start=1	
/transl_table=11	
/protein_id="AAB41867.1"	
/db_xref="GI:1814317"	
/translation="MYHSRPIPIFNIDSSNIGRRELKFLRHVEIAYRSPAOBLLIAR	
ORSLNSIINIDVPIIDPAIDEVNRPCEVACPGAPILTIILNDSLMWYCSFREM	
BRVAKRLYDVSQDSGVIVSTVPYPTGIRKETRMSQNNNGENTSDETINVDYIYF	
LPTISLQVPSHMKIRCNSEFLAPSRVDFCEIITAVGKALFEDNGLKAKPKLGMENA	
LTSIMKRLV"	
complement (3569..4972)	
/gene="1aah"	
complement (3569..4972)	
/gene="1aah"	
/codon_start=1	
/transl_table=11	
/protein_id="AAB41868.1"	
/db_xref="GI:1814318"	
/translation="MVSITSLAOSLEHLRRKQVSCLEIVETLIARCEAAKSLNALLAT	
DWDGIRRAKKIDIRGNAGVLCGIPICFRKANITGTFPISATPALINHLPIKIPSV	
AERLPISACALVGSAGNHELSFGITSNVYATGAVRWNPSSLISGSSGGAAVAARV	
LMLGSIQDTCASGRNPAALCGVVGFPPTGARVPRDRIIVPSPTRDAGIYIAOCVADV	
IILQVSGISGRARILPIPLKGLRIGLPTTYFVDLDAVDVFAAETIRLLANGGVFEV	
EADIPLEDHXXGASLPALAYEPHALKOYLDVDFVSPSDVINGIRSPDVANIYKA	
QVDEGHQSSDEVELAROSFRRLQATYRNFRLYQDAILPPTAPLAAKIACDFST	
HNGSHMNTFQIKYKRVDPSSVAGLPLVHLTPRKLPGMEIDGLAGSDHRLAIG	
AALKEALVDIHLFPVVLN"	
5306..8720	
/gene="1aah"	
join(5306..7243,8396..8720)	
/gene="1aah"	
/note="interrupted by a bacterial IS870-like insertion	
element"	
/pseudo	
/codon_start=1	
/transl_table=11	
7244..7247	
/note="Flanking repeat, left border of the IS870-like	
insertion element"	
/tpc_type=direct	
7301..8395	
/gene="ORF1"	
7301..8395	
/gene="ORF1"	
/codon_start=1	
/transl_table=11	
/protein_id="AAB41869.1"	
/db_xref="GI:1814319"	
/translation="MRITWCRTGVSFTVSTDHORLMALTRDRNAPQKHWRSEIILL	
SADQVETIIRKQIGSKTICVMRQSEFPAEGVGLRDKTRPSRIYRLSDPAVERVY	
ARRLLIRRAARQHTADMMATVAGISASAVRITWKAHGLAPHRWRQFKSNDEKFDVK	

```
LRDVGIVYDPAHAIVLSYDEKSOIQALDRTOPGLPMKKRGLGTWTHDYKRKHTTLL
PALANVDGTVICGMQRHROEPIRELFKRIINAQDPDKSVHILDNVAAHKPKYRA
MLERHFRFHTPTSCSMLNAVEGFARSLKRKRGVHVSVDLQAAINRLVBN
KKRPPTRADPDKIITAAVKGHOLDISH"
repeat_region
/nc="flanking repeat; right border of the 15870-lik
insertion element"
/rpt_type=direct
9419..10141
/gene="idp"
9419..10141
/gene="idp"
/codon_start=1
/ranal_cable=1
/protein_id="AAB41870.1"
/db_xref="gi:1814320"
/translation="MDRLRFPGPTCKGKSTAIRLAQOTLPVLSLDRVOCROSTG
SGRPVEBELKGTRLYLEDRPLVKGIIIAKQAHRLIGEVYNEAHGGLIBEGSISL
LRCAQSSVSTDFRWHIIRKLADSETFNNAKARVOMLRPAVPSIIQELVHLMN
EPLRLILKEIDGVRVAMLPASQNIPTDMLQLDPDMGELIHGIAOEYLIHAROE
QEPFVSVAVEGEPFGMC"
complement(10816..11439)
/gene="6b"
complement(10816..11439)
/gene="6b"
/codon_start=1
/ranal_cable=1
/protein_id="AAB41871.1"
/db_xref="gi:1814321"
/translation="MTVANMQVRDLTILISAGEIISQRLBEGARTDGFALLDLYFQPS
VLSGEDEYIRFSGQELVYVLRREDIARCALRNLPSNSNFGIATVAPMLNNA
HGLNOYLQRCDFGLVNYQGPHTNOFPAIMRNSCFVASFADVINSESYFVARG
NYEBAGGDDDEVDEGEAGAAATTSQIQLNLINYPVALGSCQLSA"
complement(113214..14290)
/gene="oc8"
complement(113214..14290)
/gene="oc8"
complement(113214..14290)
/pseudo
/codon_start=1
/ranal_cable=1
/14719..14745
/nc="flanking repeat; right border of the T-DNA"
/rpt_type=direct
repeat_region
BASE COUNT 3854 a 3510 c 3632 g 3958 t 6 others
ORIGIN
Query Match 69.7%; Score 1581.4; DB 1; Length 14960;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 1733; Conservative 0; Mismatches 216; Indels 2; Gaps 2;
```

```
QY 1 ATGTCAGCTCACCTCTCTTGATTAACGAGGATCATCTCCCAACCAAAATGATGAT 60
Db 5306 ATGTCAGCTCACCTCTCTTGATTAACGAGGATCATCTCTTACCAAAATGATGAT 5365
QY 61 CTGACATGTCGATTAAGCGGATGAATTGACCGCAGGGTTCCAGTCCCTCTTGAA 120
Db 5366 CTGATATGTCGATTAAGCGGATGAATTGACCGCAGGGTTCCAGTCCCTCTTGAA 5425
QY 121 CGAAGACCTTCTAGGGAAGAGATTACTCAAAAGCTTCCAGAGTCCAGCGCTGAG 180
Db 5426 CGAAGACCTTCTAGGGAAGAGATTACTCAAAAGCTTCCAGAGTCCAGCGCTGAG 5485
QY 181 GCTTGAAGGCGGCGGATGTCGCTCCCGAGATCTCAGTGGTGAAGAGTGA 240
Db 5486 GCTTGAAGGCGGCGGATGTCGCTCCCGAGATCTCAGTGGTGAAGAGTGA 5545
QY 241 GTTCTCTCCCTTATATCTATTTGCAAGAAATTCGGGCGGATCTGATCGAAA 300
Db 5546 GTCTCTCCCTTATATCTATTTGCAAGAAATTCGGGCGGATCTGATCGAAA 5605
QY 301 CTTTGGCGCGGCGCAACGATGTCGCTGCTGCTCACTGCACTTGACACCATTTGATG 360
Db 5606 CCAATGGCGCGGCGCAACGATGTCGCTGCTGCTCACTGCACTTGACACCATTTGATG 5665
```

```
QY 361 GATTTCCGAAACACAACTAAATCCAAAGCCGTGTTTCTGAGCCGTAAGATGCA 420
Db 5666 GATTTCCGAAACACAACTAAATCCAAAGCCGTGTTTCTGAGCCGTAAGATGCA 5725
QY 421 CCGATGATCTTATGATCTTCTGAGCCGTAAGATCTTCAAGCTCCGCTTCCAA 480
Db 5726 CCGAGGACCTTATGATCTTCTGAGCCGTAAGATCTTCAAGCTCCGCTTCCAA 5785
QY 481 CTGCAATGCGGCTGAGAGATGAGCAGATGAATGCGTTACCGGGTTTACATACC 540
Db 5786 CTGCAATGCGGCTGAGAGATGAGCAGATGAATGCGTTACCGGGTTTACATACC 5845
QY 541 CTGAAAGGCGCGTGCATTTTGAATGATGATGATGATGATGATGATGATGATG 600
Db 5846 CTGAAAGGCGCGTGCATTTTGAATGATGATGATGATGATGATGATGATGATG 5905
QY 601 TCGGCAAGTTCCTTTCACAACTGATGATGATGATGATGATGATGATGATGATG 660
Db 5906 TCGGCAAGTTCCTTTCACAACTGATGATGATGATGATGATGATGATGATGATG 5965
QY 661 TGTTCGATAGTGAAGGATCGGCTCTTCCGAGAGATGTTCTTAAGCGAAAGTGG 720
Db 5966 TGTTCGATAGTGAAGGATCGGCTCTTCCGAGAGATGTTCTTAAGCGAAAGTGG 6025
QY 721 GTCAATGGCGCTGGCATTTCCGAGCTGATGATGATGATGATGATGATGATGATG 780
Db 6026 GTCAATGGCGCTGGCATTTCCGAGCTGATGATGATGATGATGATGATGATGATG 6085
QY 781 GACGATGTTCAATATATGAGAGATGATGATGATGATGATGATGATGATGATG 840
Db 6086 GACGATGTTCAATATATGAGAGATGATGATGATGATGATGATGATGATGATG 6145
QY 841 TTCAGGAGGCTCTGATGTCGAGCGGAGAAAGGAGGAGATGATGATGATGATG 900
Db 6146 TTCAGGAGGCTCTGATGTCGAGCGGAGAAAGGAGGAGATGATGATGATGATG 6205
QY 901 TTCTGCTTGTGTTTCTTCTCGAGCTTACGAGCTTCTTCTGATGAGCGGTTCC 960
Db 6206 TTCTGCTTGTGTTTCTTCTCGAGCTTACGAGCTTCTTCTGATGAGCGGTTCC 6265
QY 961 CCCGCAAGTCGACACTTATCTGTCGTCACCAAGCGTCCAAATCATGTTGAAAG 1020
Db 6266 CCCGCAAGTCGACACTTATCTGTCGTCACCAAGCGTCCAAATCATGTTGAAAG 6325
QY 1021 CAGTCGACCGGAGCTTTCATTCGCTTACAGAGTGGGCTGCTTGAAGAC 1080
Db 6326 CAGTCGACCGGAGCTTTCATTCGCTTACAGAGTGGGCTGCTTGAAGAC 6385
QY 1081 GATTTTCATGAGCGAGATATGTTGTTGCTTCCGCTGCTATTACTCAGGCTTGA 1140
Db 6386 GATTTTCATGAGCGAGATATGTTGTTGCTTCCGCTGCTATTACTCAGGCTTGA 6445
QY 1141 TCAAGACATTAAGTGGGCTCATGACTCTGCGAAATTTGCTGAAACGTTTCC 1200
Db 6446 TCAAGACATTAAGTGGGCTCATGACTCTGCGAAATTTGCTGAAACGTTTCC 6505
QY 1201 GATGCTTCTCTTCAAGGATGAGAGATCTTCTGAGCAACATCTCTGAGGT-AA 1259
Db 6506 GATGCTTCTCTTCAAGGATGAGAGATCTTCTGAGCAACATCTCTGAGGT-AA 6565
QY 1260 ACATGAGTTTCTCATGATGATGAGACTTATCAAGCTAATGAGATGAGCGG 1319
Db 6566 ACATGAGTTTCTCATGATGATGAGACTTATCAAGCTAATGAGATGAGCGG -GG 6624
QY 1320 GTTGTGTCAGTTTGAAGCGGTTTATGAGATCTCTGCTTGTCTCAACGAGATA 1379
Db 6625 GTTGTGTCAGTTTGAAGCGGTTTATGAGATCTCTGCTTGTCTCAACGAGATA 6684
QY 1380 TGAAGAAATCAACGAGATGTCCTGAAGAACTCAGAACTTCCAGTGGATCGATC 1439
Db 6685 TGAAGAAATCAACGAGATGTCCTGAAGAACTCAGAACTTCCAGTGGATCGATC 6744
```

QY 1440 TGAAGTGTAAAGTGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGGCAT 1499
 Db 6745 TCAAGTGTAAAGTGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGGCAT 6804
 QY 1500 TCAAGAGAAAAGACAAAATTAAGATAGGCTTAAGAGCGGATATCTGAACCTTAATGA 1559
 Db 6805 TGAGAAAGAAAAGACAAAATTAAGATAGGCTTAAGAGCGGATATCTGAACCTTAATGA 6864
 QY 1560 TAAGGTGTGTGATCAATCTGAGCTCCGAAATATCCAACTGAGCATTTGCTGACATGCCA 1619
 Db 6865 TAAGGTGTGTGATCAATCTGAGCTCCGAAATATCCAACTGAGCATTTGCTGACATTTGA 6924
 QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGTTGATTAAGCCATATGACAGGATC 1679
 Db 6925 TACCAACATTTTTCGTACACCAAGTGAACCAAGCGTTGATTAAGCCATATGACAGGATC 6984
 QY 1680 GTCAAAACTCTTCTGTGATGACTGAACGAAAATTCGTGATGACCATATTCCTCCGTTTG 1739
 Db 6985 TTCAAAATTTTGTGTGATGACTGAACGAAAATTTTGTGATGACCATATTCCTCCGTTTG 7044
 QY 1740 TGTCTCATGAGAGGATTCGAAAAGAGTGTATGCTTGACTATGAGTGCAGATCC 1799
 Db 7045 TGTCTCATGAGAGGATTCGAAAAGAGTGTATGCTTGACTATGAGTGCAGATTC 7104
 QY 1800 GAATGTAAGTCTAGTGTCTCATCACTATATACATGAGAGAGCATCCCAAGCTGTT 1859
 Db 7105 GAACGTAAGTCTAGTGTCTCATCACTATATACATGAGAGAGCATCTCACAAGCTACT 7164
 QY 1860 GGGGGTCCCGACAAAAGAGCGATTATGCTGCGGAGCGCAATTTGAGATCTTT 1919
 Db 7165 GGGGGTCCCGACAAAAGAGCGATTATGCTGCTACGAGACGCAATTTGAAATCATTT 7224
 QY 1920 CCGGGGTTTGGCCAGACTATTTCTGCC 1950
 Db 7225 CCTGGGTTTGGCCAGACTATGTTACC 7255

Search completed: November 24, 2003, 01:37:08
 Job time : 8374.46 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 19:44:15 ; Search time 701.748 Seconds
(without alignments)
10562.608 Million cell updates/sec

Title: US-09-434-837-10
Perfect score: 2268
Sequence: 1 atgcagcttcacccctcctcct.....actacgcagtagaattag 2268

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2169961 seqs, 1634102185 residues
Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2268	100.0	2268	14	US-10-187-339-10
2	1788.6	78.9	2544	14	US-10-187-339-9
3	1788	78.8	1807	14	US-10-187-339-7
4	40.8	1.8	1049	12	US-10-140-472-358
5	40.8	1.8	1049	12	US-10-141-761-358
6	40.8	1.8	1049	12	US-10-142-885-358
7	40.8	1.8	1049	12	US-10-158-790-358
8	40.8	1.8	1049	12	US-10-137-871-358
9	40.8	1.8	1049	12	US-10-140-805-358
10	40.8	1.8	1049	12	US-10-140-864-358
11	40.8	1.8	1049	12	US-10-140-923-358
12	40.8	1.8	1049	12	US-10-141-756-358
13	40.8	1.8	1049	12	US-10-141-759-358
14	40.8	1.8	1049	14	US-10-123-155-358
15	40.8	1.8	1049	15	US-10-146-731-358
16	39.6	1.7	1695	14	US-10-156-761-6871

17	39.6	1.7	9025608	14	US-10-156-761-1	Sequence 1, Appl1
18	38.6	1.7	1575	9	US-09-912-176-2	Sequence 2, Appl1
19	37.4	1.6	7194	9	US-09-815-242-8167	Sequence 8167, Ap
20	37.4	1.6	29555	8	US-08-781-986A-206	Sequence 206, App
21	37	1.6	1007	12	US-10-027-632-10540	Sequence 10540, A
22	37	1.6	1007	12	US-10-027-632-10540	Sequence 10540, A
23	37	1.6	161652	15	US-10-081-320-40	Sequence 40, Appl
24	36.2	1.6	3758	11	US-09-948-820-42	Sequence 42, Appl
25	35	1.6	671	14	US-10-184-644-346	Sequence 346, App
26	36	1.6	671	14	US-10-184-644-346	Sequence 346, App
27	36	1.6	1578	10	US-09-880-107-3440	Sequence 3440, Ap
28	35.6	1.6	3055	12	US-10-027-632-177179	Sequence 177179,
29	35.6	1.6	3055	12	US-10-027-632-177179	Sequence 177179,
30	35.6	1.6	37332	9	US-09-873-367C-333	Sequence 333, App
31	35.2	1.6	1032	9	US-09-415-277A-4	Sequence 4, Appl1
32	35.2	1.6	1032	9	US-09-415-277A-6	Sequence 6, Appl1
33	35	1.5	368004	10	US-09-949-654-3	Sequence 3, Appl1
34	34.6	1.5	4996	12	US-10-301-822-166	Sequence 166, App
35	34.6	1.5	5068	12	US-10-301-822-164	Sequence 164, App
36	34.4	1.5	2180	14	US-10-168-274-30	Sequence 30, Appl
37	34.4	1.5	2188	14	US-10-103-312-86	Sequence 86, Appl
38	34.4	1.5	2182	12	US-10-240-965-134	Sequence 134, App
39	34.4	1.5	2291	12	US-10-320-351-16	Sequence 16, Appl
40	34.4	1.5	2291	12	US-10-320-351-17	Sequence 17, Appl
41	34.4	1.5	11745	12	US-10-240-453-205	Sequence 205, App
42	34	1.5	585	12	US-10-032-585-6692	Sequence 6692, Ap
43	34	1.5	585	12	US-10-027-632-217634	Sequence 217634,
44	34	1.5	585	13	US-10-027-632-217634	Sequence 217634,
45	34	1.5	1572	14	US-10-230-026-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-10-187-339-10
; Sequence 10, Application US/10187339
; Publication No. US2003008478A1
; GENERAL INFORMATION:
; APPLICANT: Ream, Walt et al.,
; TITLE OF INVENTION: Plants Having Enhanced Gall Resistance and
; TITLE OF INVENTION: Methods and Compositions for Producing the Same
; FILE REFERENCE: 53629
; CURRENT APPLICATION NUMBER: US/10/187,339
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US/09/434,837
; PRIOR FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 60/107,185
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
; US-10-187-339-10

Query Match 100.0%; Score 2268; DB 14; Length 2268;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ANOTGACCTTCACCTCTCTGATTAACGAGTGATCATCTCCCAACCAAAAGTGAT	60
DB	1	ATGTCACTTCACCTCTCTGATTAACGAGTGATCATCTCCCAACCAAAAGTGAT	60
QY	61	CTGACATGATGCGATTAAGCGATGAATTGACCCGAGGATTCGATGCTCTTAGAA	120
DB	61	CTGACATGATGCGATTAAGCGATGAATTGACCCGAGGATTCGATGCTCTTAGAA	120
QY	121	CGAAGGCTTTCAGGGAAGAGATTACTCAAGCTCCACCGATGACGCTGGGTTA	180
DB	121	CGAAGGCTTTCAGGGAAGAGATTACTCAAGCTCCACCGATGACGCTGGGTTA	180

QY 181 GCTTGCAAAAGCTGCGCGATGCTGCTCCCGAGATCTCAGCTGCTGGAAGGTAGCA 240
 Db 181 GCTTGCAAAAGCTGCGCGATGCTGCTCCCGAGATCTCAGCTGCTGGAAGGTAGCA 240
 QY 241 GTTCTCTCCGCTTATATCTATATTGCGCAAGAAATTCCTGGGCGGATCTTGAATCGAAA 300
 Db 241 GTTCTCTCCGCTTATATCTATATTGCGCAAGAAATTCCTGGGCGGATCTTGAATCGAAA 300
 QY 301 CTTTGGGCGGCGCAACAGTAGTGCTGCTGCTCCATCGACTTGGACACCATTTGGATG 360
 Db 301 CTTTGGGCGGCGCAACAGTAGTGCTGCTGCTCCATCGACTTGGACACCATTTGGATG 360
 QY 361 GATTTCCTCGAAGCACTAATTCAGAGCCCTGTTTGTGCTGAGCGGTAAAGATGCA 420
 Db 361 GATTTCCTCGAAGCACTAATTCAGAGCCCTGTTTGTGCTGAGCGGTAAAGATGCA 420
 QY 421 CCGATGATCTTAGTCATTTCTGCGCATTTCAATCTCTAAGACTGCGGCTTTGCAAC 480
 Db 421 CCGATGATCTTAGTCATTTCTGCGCATTTCAATCTCTAAGACTGCGGCTTTGCAAC 480
 QY 481 CTGCGCAATGCGCTGTAAGAGATGCGACGATGAATGCGTTACCGGGTTTACATACC 540
 Db 481 CTGCGCAATGCGCTGTAAGAGATGCGACGATGAATGCGTTACCGGGTTTACATACC 540
 QY 541 CTTGAAGGGGCGGCGCATTTGACATGATGCTTGAAGCTTGAAGCTGATGCTGAAGGT 600
 Db 541 CTTGAAGGGGCGGCGCATTTGACATGATGCTTGAAGCTTGAAGCTGATGCTGAAGGT 600
 QY 601 TCGGCAAGTCTCTTCCCAACATGCACTTCTCTACGATGCGACGCGTTTGTGACCA 660
 Db 601 TCGGCAAGTCTCTTCCCAACATGCACTTCTCTACGATGCGACGCGTTTGTGACCA 660
 QY 661 TGTTCGCAATGCTGACGCGATCGGCTTCTTCCGAGAGATGCTTCCGAGAGTGGCG 720
 Db 661 TGTTCGCAATGCTGACGCGATCGGCTTCTTCCGAGAGATGCTTCCGAGAGTGGCG 720
 QY 721 GTCATTTGGCGCTGCGATTTCCGAGCTGCTGCTGCGCAACGAACTGCTTCACTGCGGCTA 780
 Db 721 GTCATTTGGCGCTGCGATTTCCGAGCTGCTGCTGCGCAACGAACTGCTTCACTGCGGCTA 780
 QY 781 GACGATGTTAATATATGAGAGCAAGTATGCTGTTGAGGCAAGCTTTGCTCATGCT 840
 Db 781 GACGATGTTAATATATGAGAGCAAGTATGCTGTTGAGGCAAGCTTTGCTCATGCT 840
 QY 841 TTTCAAGGAGCGCTCTATGCTGCTGCGCAAAATGCGGCGGATGCGATTTCTCTGCTGCA 900
 Db 841 TTTCAAGGAGCGCTCTATGCTGCTGCGCAAAATGCGGCGGATGCGATTTCTCTGCTGCA 900
 QY 901 TTTGCTGCTGTTTTCTCTCTGAGCGTTACGCGCTGCTTCTGATGAGCGGTTCCAAAT 960
 Db 901 TTTGCTGCTGTTTTCTCTCTGAGCGTTACGCGCTGCTTCTGATGAGCGGTTCCAAAT 960
 QY 961 CCGGCGCAGTGCACACTTACTTGTCTACCAAGCGCTCCAAATCATGTGGAAGCGGG 1020
 Db 961 CCGGCGCAGTGCACACTTACTTGTCTACCAAGCGCTCCAAATCATGTGGAAGCGGG 1020
 QY 1021 CAGCTGCCACCGAAGCTGCTTCATCGGCTTTACAAGGCTTGGCTGCTTCTGGAAGAC 1080
 Db 1021 CAGCTGCCACCGAAGCTGCTTCATCGGCTTTACAAGGCTTGGCTGCTTCTGGAAGAC 1080
 QY 1081 GGTTCATAGGAGATATTTGTTGCTTGGCTTGGCTGCTGCTATTAATCAAGCGCTTGA 1140
 Db 1081 GGTTCATAGGAGATATTTGTTGCTTGGCTTGGCTGCTGCTATTAATCAAGCGCTTGA 1140
 QY 1141 TCAGGACACATTAAGTGGCTCATGCTGCGCAATTTGGCTGGAAGCGTTTGGAGG 1200
 Db 1141 TCAGGACACATTAAGTGGCTCATGCTGCGCAATTTGGCTGGAAGCGTTTGGAGG 1200
 QY 1201 GAGTCTTCTCTTCAAGGATATGAGAGATCTTCTGCGCAACATCTCTGCTGCTGTA 1260
 Db 1201 GAGTCTTCTCTTCAAGGATATGAGAGATCTTCTGCGCAACATCTCTGCTGCTGTA 1260
 QY 1261 CATGAGATTTTCTCTCATGATTTGGGACCTATTAAGCTAATGAGATCTGCGGGG 1320

Db 1261 CATGAGATTTTCTCTCATGATTTGGGACCTATTAAGCTAATGAGATCTGCGGGG 1320
 QY 1321 TTTGCTCAGATTTTGAAGCGGTTTATGAATCTCTGCTTGGCTATCAACGATAT 1380
 Db 1321 TTTGCTCAGATTTTGAAGCGGTTTATGAATCTCTGCTTGGCTATCAACGATAT 1380
 QY 1381 GAAGAAATCAGGAGATGCTGCTGAGGAAATCTCAGAACTTCAAGCTGCGATCT 1440
 Db 1381 GAAGAAATCAGGAGATGCTGCTGAGGAAATCTCAGAACTTCAAGCTGCGATCT 1440
 QY 1441 GAAGGTTAACGCTGCTGCTGAGCAGCGGATATGCGATCTTCAAGTCAAGGCGAT 1500
 Db 1441 GAAGGTTAACGCTGCTGCTGAGCAGCGGATATGCGATCTTCAAGTCAAGGCGAT 1500
 QY 1501 CAGAGGAAAGCAAAATTAAGATTAAGCTTAAAGCGGATATCTGAACCTTATGAT 1560
 Db 1501 CAGAGGAAAGCAAAATTAAGATTAAGCTTAAAGCGGATATCTGAACCTTATGAT 1560
 QY 1561 AAGTGTGTGCTCATCTGAGCTGCGAAATATCCAATCAGGATTTGCTGACATGCGAT 1620
 Db 1561 AAGTGTGTGCTCATCTGAGCTGCGAAATATCCAATCAGGATTTGCTGACATGCGAT 1620
 QY 1621 ACCAATATTTTTCAGGACCGAGTGAACCAAGCGTTGATTAAGCCATATGACAGATCG 1680
 Db 1621 ACCAATATTTTTCAGGACCGAGTGAACCAAGCGTTGATTAAGCCATATGACAGATCG 1680
 QY 1681 TCAAAATCTCTCTGATGACTGAACGAAATTTCTGTTAACAATCTCTCCGCTTGT 1740
 Db 1681 TCAAAATCTCTCTGATGACTGAACGAAATTTCTGTTAACAATCTCTCCGCTTGT 1740
 QY 1741 GTCCCTATGAGAGGAGATGCGAAAGCAGTATTTGCTGAGCTATGAGTGCAGAGATCCG 1800
 Db 1741 GTCCCTATGAGAGGAGATGCGAAAGCAGTATTTGCTGAGCTATGAGTGCAGAGATCCG 1800
 QY 1801 AATGTAAGTCTAGTCTCATCAGTTATATATGAGAGCACTTCCCAAGCTGTTG 1860
 Db 1801 AATGTAAGTCTAGTCTCATCAGTTATATATGAGAGCACTTCCCAAGCTGTTG 1860
 QY 1861 GCGGTCCCGCAAAAAGAGCATTAATGCTGCTGCGGACCGCAATTTGAGATCTTTC 1920
 Db 1861 GCGGTCCCGCAAAAAGAGCATTAATGCTGCTGCGGACCGCAATTTGAGATCTTTC 1920
 QY 1921 CCGGCGTTGCGCAGACCTAATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATCAA 1980
 Db 1921 CCGGCGTTGCGCAGACCTAATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATCAA 1980
 QY 1981 CATGATTTGCTTACAGAGCAATGCGGCGGAGCTTCAAACTCAACCGGCTGCTGAG 2040
 Db 1981 CATGATTTGCTTACAGAGCAATGCGGCGGAGCTTCAAACTCAACCGGCTGCTGAG 2040
 QY 2041 GATTTTATTTCTGAAGAACTTTCTTCAAGCATGCAACGCTATATACCGGAT 2100
 Db 2041 GATTTTATTTCTGAAGAACTTTCTTCAAGCATGCAACGCTATATACCGGAT 2100
 QY 2101 TACTTGGCGGTTGAGCTTGTCTTCAAGGAGATGAGTGGTGGAGGCTATTAAGACC 2160
 Db 2101 TACTTGGCGGTTGAGCTTGTCTTCAAGGAGATGAGTGGTGGAGGCTATTAAGACC 2160
 QY 2161 GCGGTAAACGCGCTCTGTCATTAATCAAAATGTTGAGGCAATTTGGCAAGGCGAAT 2220
 Db 2161 GCGGTAAACGCGCTCTGTCATTAATCAAAATGTTGAGGCAATTTGGCAAGGCGAAT 2220
 QY 2221 CTTCTGAAACACTTTGGAAGAGATTAATCAACCGCATGTAATTTG 2268
 Db 2221 CTTCTGAAACACTTTGGAAGAGATTAATCAACCGCATGTAATTTG 2268

RESULT 2
 US-10-187-339-9
 ; Sequence 9, Application US/10187339
 ; Publication No. US20030084478A1
 ; GENERAL INFORMATION:


```

US-10-140-472-358
; Sequence 358, Application US/10140472
; Publication No. US2003013888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRP
; ORGANISM: Homo Sapien
US-10-140-472-358

Query Match      1.8%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.028;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAAGGGCCGTCGATTTGACATGCTTATGCTGGAACCTGATGCTGAAG 597
DB 298 HNSLSQHVPRPFKXINKLQELDLSONFLAKEIGDKFLHFLPSLIQDLDSNFELOYVR 357
QY 598 GGTTCGACAGGTTCCTTCCCAACATCGACTGCTGACGACGACGCTTTTTCAG 657
DB 358 ASMNLSQAFFSLKSLKILIRGVFKELKSPNLSPLHNLQNLVLDLGNFIKIANLSWF 417
QY 658 CAATGTTCCGATAGTGAAGCATGCGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
DB 418 KQFKRLKVIDLSVKNISPSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKVARSCHF 477
QY 718 GCGGTATTTGGCGCTGCGCATTTCCGACATCGTGTGGCAACGACATGCTTATGCTGGG 777
DB 478 KKKEASFMVSNECSYRGQTLDSLKNSIFPVKSSDPQHSFLKCLNLSNLSQTLNGSE 537
QY 778 GTAGACGATTTAATATATATGAAGCATGATCGTTGGAGGAGCGTTGGTCAAT 837
DB 538 FOPLAELRYLDFSNRDLHLHSTAFEBLHLEVLDISSNSHYFQSEGITHMLNFTKNLKV 597
QY 838 GGTTCGACAGGCTCTCTAGTGTGTCGCGCAATGGGGCGATGCGATTTCTCTGCTGCT 897
DB 598 LQKLMNDNDISSSTRTWESSELRFLFRGNHLDVLMBEGDNRVYLQFLKNLKLKEBLDI 657
QY 657 GCATTTGCTGTTTTTTCTTCTCGACGCTTACGCGCTGCTTTCGATGAGCGGCTTCCA 957
DB 658 SKNSISFLPSGVFDGMPNKLKNSLAKNGLKFSWPKLOCLKNLETLDSLHNOQLTVPER 717
QY 958 AATCCGACAGTCACTACTGCTGCTACCAAGGCGTCAATCAATGGAAGCC 1017
DB 718 LSNCSLSKLNLLKNNQISLTKYFLQDAPQLRYLDSLNSKTIOMIOKTFPENVLNLMK 777
QY 1018 GGGACAGCTCCCAAGAGCTTTCATCGCGTTTACCAACGTTGGCGTCTTGAAG 1077
DB 778 LLLHNRFLCTGDAVFWVWVNHTEVTIPLYLATDVCVGBAHKQGSVISLDTYTELDL 837

```

```

QY 1078 GACGGTTTCATGAGAGATATTTGCTGCTGCTGCTGCTTACTGAGCGCTT 1137
DB 838 TWLILFSLISVSFLMVMWMTASHLYFMDVWYTHHCXAKIKGYORLISPPCCYDAFIY 897
QY 1138 AATCAGACATGATGAGGCTCATGACTCTGCGCAATTTGGCTGCAACGCTTTCGGG 1197
DB 898 DTKDPAVTEWVLAELVAKLEDREKHFNLCLEBOMLPQPVLENLSQSISKKTIVFM 957
QY 1198 AGGAGCTCTTCTCTTCGAGGATAGAGAGATCTTTCGGGACACAT 1245
DB 958 TDKYAKTENFKIAFYLSHORLMDKEDVDVLIILFLKPKFQSKFLQRLK 1005

RESULT 5
US-10-141-761-358
; Sequence 358, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRP
; ORGANISM: Homo Sapien
US-10-141-761-358

Query Match      1.8%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.028;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAAGGGCCGTCGATTTGACATGAGCTTATGCTGGAACCTGATGCTGAAG 597
DB 298 HNSLSQHVPRPFKXINKLQELDLSONFLAKEIGAKFLHFLPSLIQDLDSNFELOYVR 357
QY 598 GGTTCGACAGGTTCCTTCCCAACATCGACTGCTGACGACGACGCTTTTTCAG 657
DB 358 ASMNLSQAFFSLKSLKILIRGVFKELKSPNLSPLHNLQNLVLDLGNFIKIANLSWF 417
QY 418 KQFKRLKVIDLSVKNISPSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKVARSCHF 477
QY 478 KKKEASFMVSNECSYRGQTLDSLKNSIFPVKSSDPQHSFLKCLNLSNLSQTLNGSE 537
DB 538 GTAGACGATTTAATATATGAAGCATGATCGTTGGAGGAGCGTTGGTCAAT 837
DB 538 FOPLAELRYLDFSNRDLHLHSTAFEBLHLEVLDISSNSHYFQSEGITHMLNFTKNLKV 597

```

QY 838 GCTTCAGGAGCGCTCTAGTGTGTCGGCGAAATGGGGCGATGGCTTCTCTGCT 897
DB 598 LQKMMNDNDISSSTSRTHSESLRTLEFRGNHLDVMBGDNRYQLFRNLKLELDI 657
QY 898 GCATTCGCTGTTTCTTCTCGAGCGCTTACGCGCTGCTCGATGAGCGCTTCCA 957
DB 658 SKNSLSFLPSGVFDGMPNPKULSLAKNGLKFSWKKLOCLKNLETLDSHNLTVPER 717
QY 958 AATCCCGGACAGTCACACTTACTTGTCCTACCAAGCGCTCCAAATACATGTGGAAGCC 1017
DB 718 LNSCSRSLKNLILKNNQIRSLTKYFLQDAFOLRYLDSLNSNKIOMIKTSPENVNLNLM 777
QY 1018 GGGCAGCTGCGACCGAGCGTTCATCGGCTTACACAGCTGGCGCTGCTTGAAG 1077
DB 778 LLIHNRFLCTCDVAFVWVNVNTEVTIPLATDVTCVGRNKGQSVISLDTYCELD 837
QY 1078 GACGCTTTTCATGAGAGATATTGTGTGGCTTGGCTTGGCTTCTTACTTACGCGCTTG 1137
DB 838 TNLILFSLISVSLFLMWMNTASHLYFWVDVWYIYHFCCKAKIKGYORLISPDCCYDAFIY 897
QY 1138 AATCAGACACATTAGTGGGCTCATGACTCTGGCAATTGGCTGAACGCTTCCGG 1197
DB 898 DTKDPATVEWVLAELVAKLEDPREKHFNLCLBERDWLPGQVLENTSOSIQLSKKTIVFWM 957
QY 1198 AGGAGCTCTCTCTCAGGAGATAGAGAGATCTTCTGGGACACAT 1245
DB 958 TDKYAKTENFKIAFYLSHORLMDKVDVILLIFLEKPFQSKFLQLRK 1005

RESULT 6

US-10-142-885-358
; Sequence 358, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Matanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C248
; CURRENT APPLICATION NUMBER: US/10/142,885
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-885-358

Query Match 1.8%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.028;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;
QY 538 ACCCTGAAGGGCGCGCCATTGACATGTGCTTATGCTGAACCTATGCTGAAG 597
DB 298 HSNLSQHVPPWFNKNKLELDLSONFLAKEIDAKLHFLPSLIQDLSFNFELQYR 357
QY 598 GGTTCGGACGCTCTCTTCCAAACATGACTTGTGCTACGACTGACACGCTTTTGGAC 657

DB 358 ASMNLSQAFSSLSIKILIRIRGVFEKLSFNLSPHLNLDNEVLDTGNFIKIANLSMF 417
QY 658 CATGTTCGCAATGAGACAGCATCGGCTCTTTCGGAGAGATGTTCTTAAGCGGAAGTG 717
DB 418 KQFKULYDLISVNIISPESDSSEWFGSMARTSVESIEPVOLHYRIRYKVASCF 477
QY 718 GCGTCATTCGCGCTGCAATTCGGAATCTGTCGGCGCAAGCACTGCTTACGCTGGG 777
DB 478 KXKESFMSVNESCYKGTGLDLSKNSIFVASSPDQHSFLKCNLSGNLISQTLNGSE 537
QY 778 GTAGACATGTTACATATATAGACAGTATCGTGTGAGGCAAGCTTGGTACAT 837
DB 538 POPLAELRYLDPSNNRLDLHSTAFEEHLKLEVLDSLNSHVSFOSGITHMNFNKLKY 597
QY 838 GCTTCAGGAGCGCTCTGTCGTGCGCGCGAAATGGGGCGATGGATTTCTCTGCT 897
DB 598 LQKMMNDNDISSSTSRTHSESLRTLEFRGNHLDVMBGDNRYQLFRNLKLELDI 657
QY 898 GCATTCGCTGTTTCTTCTCGAGCGCTTACGCGCTGCTTTCGATGAGCGCTTCCA 957
DB 658 SKNSLSFLPSGVFDGMPNPKULSLAKNGLKFSWKKLOCLKNLETLDSHNLTVPER 717
QY 958 AATCCCGGACAGTCACACTTACTTGTCCTACCAAGCGCTCCAAATACATGTGGAAGCC 1017
DB 718 LNSCSRSLKNLILKNNQIRSLTKYFLQDAFOLRYLDSLNSNKIOMIKTSPENVNLNLM 777
QY 1018 GGGCAGCTGCGACCGAGCGTTCATCGGCTTACACAGCTGGCGCTGCTTGAAG 1077
DB 778 LLIHNRFLCTCDVAFVWVNVNTEVTIPLATDVTCVGRNKGQSVISLDTYCELD 837
QY 1078 GACGCTTTTCATGAGAGATATTGTGTGGCTTGGCTTGGCTTCTTACTTACGCGCTTG 1137
DB 838 TNLILFSLISVSLFLMWMNTASHLYFWVDVWYIYHFCCKAKIKGYORLISPDCCYDAFIY 897
QY 1138 AATCAGACACATTAGTGGGCTCATGACTCTGGCAATTGGCTGAACGCTTCCGG 1197
DB 898 DTKDPATVEWVLAELVAKLEDPREKHFNLCLBERDWLPGQVLENTSOSIQLSKKTIVFWM 957
QY 1198 AGGAGCTCTCTCTCAGGAGATAGAGAGATCTTCTGGGACACAT 1245
DB 958 TDKYAKTENFKIAFYLSHORLMDKVDVILLIFLEKPFQSKFLQLRK 1005

RESULT 7

US-10-158-790-358
; Sequence 358, Application US/10158790
; Publication No. US20030180879A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Matanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C448
; CURRENT APPLICATION NUMBER: US/10/158,790
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550

SEQ ID NO 358
 LENGTH: 1049
 TYPE: PRF
 ORGANISM: Homo Sapien
 US-10-158-790-358

Query Match 1.8%; Score 40.8; DB 12; Length 1049;
 Best Local Similarity 5.4%; Pred. No. 0.028;
 Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTGAAGGGCCGCGCATTTGACATGATGCTTATGCTGAAACCTGATGCTGAAG 597
 DB 298 HSNLSLHVPFRPFKNINKLOEIDLSQNFLEKIDGAKFHLPBLSLQDLDFNFELOVR 357
 QY 598 GGTTCGAGAGTTCCTTCCACATGACATGCTGCTGACAGACCGCTTTTGGAC 657
 DB 358 ASNMISQAFSSLSKLIRGVFPELKSFNLSPLHNLONLEVDLGNFTKIANLSMF 417
 QY 658 CAATGTTCCGATAGTGCAGCATGCGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
 DB 418 KQFKRLKVIDLSVKNKISPSGDSSEVFCNSNARTSVESYEPVLEQLHYFRYDKYARSCR 477
 QY 718 GCGGATATGGCGCTGCGCATTTCCGACATGCTGCTGCGCAAGCAACTGCTTATGCTGG 777
 DB 478 KXKEASFMVSNECYKYGQTLDSLKNSIFPVKSSDFQHLFSLKCLNLSGNLSQTLNGSE 537
 QY 778 GTAGAGATGTTCAATATATGAAGCAAGATCGCTTGTGAGGCAAGCTTGTGCAT 837
 DB 538 FQPLAELRYDLFSNNRDLHSHAFEBELHLEVDLSSNSHYFQSEGITTHMLNFTKLV 597
 QY 838 GCTTTCAGGAGCGCTCTTCTTCCGACATGCTGCGCAAGTGGGCGCATGCTTCTGCT 897
 DB 598 LQKLMMNDNDISSSTRTMESESLRTLEFRGNHLDVLMREGDNRVYLQFLNKLKLELDI 657
 QY 898 GCAATTCGCTGTTTCTTCTTCCGACGCTTACGCGCTGCTTGTGAGGCGCTTCCA 957
 DB 658 SKNSLSFLPSGVFDGPNPKNLISLAKNGLKFSWKLOCLKNLETLDSLHNOITTVPER 717
 QY 958 AATCCCGGACAGTCCACACTTACTGCTTACCAAGCGCTCCAAATACATGTGAAAGCC 1017
 DB 718 LSNCSLSLKNLILKNNQIRSLTYFLQDAFQRLRYLDLSSNKMIOIKTSPENVNLNLM 777
 QY 1018 GGCAGAGTCCGACGAGCTGTTCAATGCTTACCAAGCGCTGCGCTTGTGAAG 1077
 DB 778 LLLHNRFLCTCDVAVFVWVNHTEVTIPLYLATDVTCVGRPAKGSVISLDYTELDI 837
 QY 1078 GAGCGTTTCATGAGGAGATATGTTGCTGCTGCTGCTGCTGCTTACTTACAGGCTTG 1137
 DB 838 TNLILFSLISVSLFLMVMWMTASHLYFWDVWYIYHFCAKIKGYORLISPDCCYDAFIY 897
 QY 1138 AATTCAGACACATTAGTGGGCTCATGACTCCGCGCAATTTGGCTGAACCGTTTCGG 1197
 DB 898 DTKDAVTEWVLAELVAKLEDPREKFNLCLEBRDMLPGQPVLENLSQISLTKTVFW 957
 QY 1198 AGGAGTCTTCTTCTTCCGAGATAGAGAGATCTTTCGCGCACACAT 1245
 DB 958 TDKXAKTENFKIAFYLSHQRLMDEKVDVILLIFLEKFPQSKFLQLRK 1005

RESULT 8
 US-10-137-871-358
 Sequence 358, Application US/10137871
 Publication No. US20030207350A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Beresini, Maureen
 APPLICANT: DeForge, Laura
 APPLICANT: Desnoyers, Luc
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Geriktsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.

APPLICANT: Guiney, Austin L.
 APPLICANT: Sherwood, Steven
 APPLICANT: Smith, Victoria
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K
 APPLICANT: Wood, William
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P330R1C153
 CURRENT APPLICATION NUMBER: US/10/137,871
 PRIOR APPLICATION DATE: 2002-05-03
 NUMBER OF SEQUENCE ID NOS: 550
 SEQ ID NO 358
 LENGTH: 1049
 TYPE: PRF
 ORGANISM: Homo Sapien
 US-10-137-871-358

Query Match 1.8%; Score 40.8; DB 12; Length 1049;
 Best Local Similarity 5.4%; Pred. No. 0.028;
 Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTGAAGGGCCGCGCATTTGACATGATGCTTATGCTGAAACCTGATGCTGAAG 597
 DB 298 HSNLSLHVPFRPFKNINKLOEIDLSQNFLEKIDGAKFHLPBLSLQDLDFNFELOVR 357
 QY 598 GGTTCGAGAGTTCCTTCCACATGACATGCTGCTTACAGTGCAGACCGCTTTTGGAC 657
 DB 358 ASNMISQAFSSLSKLIRGVFPELKSFNLSPLHNLONLEVDLGNFTKIANLSMF 417
 QY 658 CAATGTTCCGATAGTGCAGCATGCGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
 DB 418 KQFKRLKVIDLSVKNKISPSGDSSEVFCNSNARTSVESYEPVLEQLHYFRYDKYARSCR 477
 QY 418 GCGGATATGGCGCTGCGCATTTCCGACATGCTGCTGCGCAAGCAACTGCTTATGCTGG 777
 DB 478 KXKEASFMVSNECYKYGQTLDSLKNSIFPVKSSDFQHLFSLKCLNLSGNLSQTLNGSE 537
 QY 778 GTAGAGATGTTCAATATATGAAGCAAGATCGCTTGTGAGGCAAGCTTGTGCAT 837
 DB 538 FQPLAELRYDLFSNNRDLHSHAFEBELHLEVDLSSNSHYFQSEGITTHMLNFTKLV 597
 QY 838 GCTTTCAGGAGCGCTCTTCTTCCGACATGCTGCGCAAGTGGGCGCATGCTTCTGCT 897
 DB 598 LQKLMMNDNDISSSTRTMESESLRTLEFRGNHLDVLMREGDNRVYLQFLNKLKLELDI 657
 QY 898 GCAATTCGCTGTTTCTTCTTCCGACGCTTACGCGCTGCTTGTGAGGCGCTTCCA 957
 DB 658 SKNSLSFLPSGVFDGPNPKNLISLAKNGLKFSWKLOCLKNLETLDSLHNOITTVPER 717
 QY 958 AATCCCGGACAGTCCACACTTACTGCTTACCAAGCGCTCCAAATACATGTGAAAGCC 1017
 DB 718 LSNCSLSLKNLILKNNQIRSLTYFLQDAFQRLRYLDLSSNKMIOIKTSPENVNLNLM 777
 QY 1018 GGCAGAGTCCGACGAGCTGTTCAATGCTTACCAAGCGCTGCGCTTGTGAAG 1077
 DB 778 LLLHNRFLCTCDVAVFVWVNHTEVTIPLYLATDVTCVGRPAKGSVISLDYTELDI 837
 QY 1078 GAGCGTTTCATGAGGAGATATGTTGCTGCTGCTGCTGCTGCTTACTTACAGGCTTG 1137
 DB 838 TNLILFSLISVSLFLMVMWMTASHLYFWDVWYIYHFCAKIKGYORLISPDCCYDAFIY 897
 QY 1138 AATTCAGACACATTAGTGGGCTCATGACTCCGCGCAATTTGGCTGAACCGTTTCGG 1197
 DB 898 DTKDAVTEWVLAELVAKLEDPREKFNLCLEBRDMLPGQPVLENLSQISLTKTVFW 957
 QY 1198 AGGAGTCTTCTTCTTCCGAGATAGAGAGATCTTTCGCGCACACAT 1245
 DB 958 TDKXAKTENFKIAFYLSHQRLMDEKVDVILLIFLEKFPQSKFLQLRK 1005


```
RESULT 9
US-10-140-805-358
; Sequence 358, Application US/10140805
; Publication No. US20030207417A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C176
; CURRENT APPLICATION NUMBER: US/10/140,805
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-805-358

Query Match      1.8%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.028;
Matches 38; Conservativity 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAAGGGCGCGCATTTGACATGATGCTTATGCTGAAACCTGATGCTGAAG 597
DB 298 HNSLSLGHVPPRWKINIKLOELDLSONFLAKEIGDAKFLHFLPSLIQLDLSFNFELQYVR 357
QY 598 GGTTCGACAGCTTCCTTCCAAACATGACATGCTCTACAGACGACGACGCTTTTGGAC 657
DB 358 ASNNLSQAFSSLSKLIRIGVFKELSFNLSPLNINONLEVLDTGTFIKIANLSMF 417
QY 658 CAATGTTCCGATAGTGACGAGATCGCTCTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
DB 418 KQPKRLKVIDLSVNKISPSGDSSEVFCSNARTSVSEYEOVLEQLHFFRYDKYARSCR 477
QY 718 GCGGTCAATGGCGCTGACATTTCCGACATCGTGTGTCGCAACGAACGATGCTTACGCTGG 777
DB 478 KKEASPMVSNECYKYGQTLDSKNSIFPVKSDPHLSFLKLNLSGNLISQTLNGSE 537
QY 778 GTAGACATGTTCAATATATGAAGCAAGTATGCTTGGAGGCAAGCTTTGGTCACAT 837
DB 538 FQPLAEIRYIDFSNNRDLHSTAFEBLHLEVLDISNSHYROSEGITTMNLNFTKNLV 597
QY 838 GCTTTCGAGAGCCTCTAGTGTGTCGTCGCAAAATGGGGCGAGATGCTTCTCTCTGCT 897
DB 598 LQTLMMNDNISSTSTSTSESRLETFEFGNHLVDVIMREGDNRYLDLFFKNLKLBLELDI 657
QY 898 GCAATTCGCTGTTGTTTCTTCCGACGCTTACGCGCTGTCTTTCGATGAGCGCGCTTCCA 957
DB 658 SKNSLSFLPSGVEFGMPNKNLSLAKNGLKSFSWKLQCLKNLETLDLSHNQLTTPVER 717
QY 958 AATCCCGGACAGTCGACATTAAGTCTTACCAAGCGCTCCAAATCATGTGGAAGCC 1017
DB 718 LSNCSRLKNLILKNNOIRSLTKFYLDAPQLRLDLSNNKIQITQSPENLVNLNLM 777
QY 1018 GGACAGCTGCCACGAAGCTGTTCCATGCGGTTTACAACGATTGGCGCTGCTTGAAG 1077
```

```
US-10-140-864-358
; Sequence 358, Application US/10140864
; Publication No. US20030207419A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C184
; CURRENT APPLICATION NUMBER: US/10/140,864
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper
; NUMBER OF SEQ ID NOS: 550
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-864-358

Query Match      1.8%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.028;
Matches 38; Conservativity 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAAGGGCGCGCATTTGACATGATGCTTATGCTGAAACCTGATGCTGAAG 597
DB 298 HNSLSLGHVPPRWKINIKLOELDLSONFLAKEIGDAKFLHFLPSLIQLDLSFNFELQYVR 357
QY 598 GGTTCGACAGCTTCCTTCCAAACATGACATGCTCTTACGACATGACGACGCTTTTGGAC 657
DB 358 ASNNLSQAFSSLSKLIRIGVFKELSFNLSPLNINONLEVLDTGTFIKIANLSMF 417
QY 658 CAATGTTCCGATAGTGACGAGATCGCTCTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
DB 418 KQPKRLKVIDLSVNKISPSGDSSEVFCSNARTSVSEYEOVLEQLHFFRYDKYARSCR 477
QY 718 GCGGTCAATGGCGCTGACATTTCCGACATCGTGTGTCGCAACGAACGATGCTTACGCTGG 777
DB 478 KKEASPMVSNECYKYGQTLDSKNSIFPVKSDPHLSFLKLNLSGNLISQTLNGSE 537
QY 778 GTAGACATGTTCAATATATGAAGCAAGTATGCTTGGAGGCAAGCTTTGGTCACAT 837
```


Db 538 FOPLAELRYLDFSNRNLDLHSTAEBELHLEVLIDISSNSHYFQSEGITMHLFTKLV 597
QY 838 GCTTTCAGGAGGAGCTCTAGTGTGTCGGCGGAATGGGGCGATTCCTCTCTGCT 897
Db 598 LQKLMNNDNDISSSTRWSESRLTFEFGNHLVDLMREGDNRVYQLFKNLKLLELDI 657
QY 898 GCATTCGCTGTTGTTTCTCTCTCGAGCGCTTACGGCGCTGCTCGATGAGCGCTTCCCA 957
Db 658 SKNSLSFLPSGVDPMPNLKXLSLAKNGLKSFWSKLOCLKNLETLDSLHNDLTVPER 717
QY 958 AATCCCGGACAGTGCACACTTACTTGCTTACCAAGCGCTCCAAATACATGTAAGAACCC 1017
Db 718 LSNCSRLKXNLILKNNQIRSLTKYFLQDAPQLRYLDSLNSKIQIKTSPFENVNLNLM 777
QY 1018 GGGCAGCTGCGACCGAGCTTCCATCGGCTTACACGGCTTGGCGCTGCTTGAAG 1077
Db 778 LLIHNRFLCTCDVAFVFWVWVNHTEVTIPLYLATDVTCVPGAHKGSVSLDLYTELDL 837
QY 1078 GAGCGTTTCATGAGAGATATGTTGTTGGCTTCCGCTGCTGCTATTACTGAGCGCTTG 1137
Db 838 TNLILFSLISVSLFLMVMWMTASHLYFWDVWYIYHCKAKIKGYORLISPDCCYDAFIY 897
QY 1138 AATCAGGACACTTATGAGTGGGCTCATGACTCTCGCAATTGGCTGAACCGTTTCGGG 1197
Db 898 DTKDPATVEWVLAELVAKLEDPREKHFNLCLEBDMLPGQVLENLSSQISQSKTVFWM 957
QY 1198 AGGAGTCTCTCTCTTCAGGAGATAGAGAGATCTTCTGGGACACAT 1245
Db 958 TDKYAKTENFKIAFYLSHORLMDKVDVILIFLEKPFQSKFLQLRK 1005

RESULT 11

US-10-140-923-358
; Sequence 358, Application US/10140923
; Publication No. US20030207355A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C188
; CURRENT APPLICATION NUMBER: US/10/140, 923
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-923-358

Query Match 1.8%; Score 40.8; DB 12; Length 1049;

Best Local Similarity 5.4%; Pred. No. 0.028; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAAGGGGCGCTGCAATTGACATGCTTATGTCGAACCTGATGCTGAAG 597
Db 298 HSNLSGHVPRWPKNKINKQELDLSONFLAKEIGDAKFLHFLPSLSLQDLSFNFELQYVR 357

QY 598 GGTTCGAGAGTTCCTTCCCAATCGACTGTCTTACAGCTGACGAGCGTTTTCGAC 657
Db 358 ASMNLSQAFSSLSKSLKILIRIGVFEKLSFNLSPILAHQNLBEVLDTGTFIKIANLSMF 417
QY 658 CATGTTCCGATAGTGAAGATGCGATTCCTTCCGAGAGATGTTCTTAAGCGGAAGTG 717
Db 418 KQFKRLKVLDSLWNKISPGDSSVDFGCMNATSVESYEPQVLEQHYRYDKVANSCHF 477
QY 718 GGGTCATTGGCGCTGCGACTTTCGCGACTGCTGGTGGCAAGCAAGTCTTACGCTGGG 777
Db 478 KXKESFMSVNSVCYVCGTDLDSLKNSIFVXSDFQHLFLKCLNLGNLSQTLNGSE 537
QY 778 GTAGAGATGTACATATATAGAGCAAGTATGCTGTGAGGACGAGCTTGGCTACAT 837
Db 538 FOPLAELRYLDFSNRNLDLHSTAEBELHLEVLIDISSNSHYFQSEGITMHLFTKLV 597
QY 838 GCTTTCAGGAGGAGCTCTATGTCGTGCGCGGAATGGGGCGATGCGATTCCTCTGCT 897
Db 598 LQKLMNNDNDISSSTRWSESRLTFEFGNHLVDLMREGDNRVYQLFKNLKLLELDI 657
QY 898 GCATTCGCTGTTTCTCTCTCGAGCGCTTACGGCGCTGCTCGATGAGCGCTTCCCA 957
Db 658 SKNSLSFLPSGVDPMPNLKXLSLAKNGLKSFWSKLOCLKNLETLDSLHNDLTVPER 717
QY 958 AATCCCGGACAGTGCACACTTACTTGCTTACCAAGCGCTCCAAATACATGTAAGAACCC 1017
Db 718 LSNCSRLKXNLILKNNQIRSLTKYFLQDAPQLRYLDSLNSKIQIKTSPFENVNLNLM 777
QY 1018 GGGCAGCTGCGACCGAGCTTCCATCGGCTTACACGGCTTGGCGCTGCTTGAAG 1077
Db 778 LLIHNRFLCTCDVAFVFWVWVNHTEVTIPLYLATDVTCVPGAHKGSVSLDLYTELDL 837
QY 1078 GAGCGTTTCATGAGAGATATGTTGTTGGCTTCCGCTGCTGCTATTACTGAGCGCTTG 1137
Db 838 TNLILFSLISVSLFLMVMWMTASHLYFWDVWYIYHCKAKIKGYORLISPDCCYDAFIY 897
QY 1138 AATCAGGACACTTATGAGTGGGCTCATGACTCTCGCAATTGGCTGAACCGTTTCGGG 1197
Db 898 DTKDPATVEWVLAELVAKLEDPREKHFNLCLEBDMLPGQVLENLSSQISQSKTVFWM 957
QY 1198 AGGAGTCTCTCTCTTCAGGAGATAGAGAGATCTTCTGGGACACAT 1245
Db 958 TDKYAKTENFKIAFYLSHORLMDKVDVILIFLEKPFQSKFLQLRK 1005

RESULT 12

US-10-141-756-358
; Sequence 358, Application US/10141756
; Publication No. US20030207359A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C200
; CURRENT APPLICATION NUMBER: US/10/141, 756
; CURRENT FILING DATE: 2002-05-08

; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 358
 ; LENGTH: 1049
 ; TYPE: PR1
 ; ORGANISM: Homo Sapien
 US-10-141-756-358

Query Match 1.8%; Score 40.8; DB 12; Length 1049;
 Best Local Similarity 5.4%; Pred. No. 0.028;
 Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAAGGCGCCGTCATTGACATGTAAGTGTGCGAAACCTGATGCTGAAG 597
 DB 298 HNSLSLGHVPRWRKINKLOEIDLSNPLAKEIDAKFLHFLPSLJQLDLSFNELOVR 357
 QY 598 GGTTCGGAGGTCCTTCCCAACATGCACTGCTTACAGATGAGACCGTTTGTGAC 657
 DB 358 ASNNLSQAFSSLKSLKILIRGVFKELKSFNLSPLHNLONLEVDLGTNFIKIANLSMF 417
 QY 658 CAATGTTCCATAGTGAAGCATGCGCTTCTTCCGAGAGATGTTCTTAAGCGGAAGTG 717
 DB 418 KQFRILKVIDLSVNKISPSGDSSEVGFCSNARTSVSEYEPQVLEQLHYFRYDKYASCRF 477
 QY 718 GCGGTCAATGCGCGTCATTTCCGAGCTCGTGGTGCAAGCAAGCTTTCATGCTGGG 777
 DB 478 KKEASMSVNESEYKQGLDLSKNSIFVKSDDPHLSFKLNSGNLISQTLNGSE 537
 QY 778 GTAGACAGTTCATATATATGAAGCAAGTATGATGCTGTTGAGCAAGCTTGTGCAT 837
 DB 538 FQPLAELRYLDFSNRDLHSTAFELHKLVDLSSNSHYQSEGITMMLNFTKLV 597
 QY 838 GCTTTCAGGAGCGCTCTAGTGTGCGCGAAATGGGGCGAGTTCCTCCGCT 897
 DB 598 LQKLMMNDNDISSSTRMESESLRLEFRGNHLDVLMREGDNRYQLDFNLLKLELDI 657
 QY 898 GCATTCGCTGTTTCTTCTCGAGCGCTTACGCGCTGCTTCGATGAGCGCGTCCA 957
 DB 658 SKNSLSLPSGVDPMPNLKNSLAKNGLSKSPKLOCLKYLETDLDSHNLTVPER 717
 QY 958 AATCCCGGACAGTGCAGACTTACTGTTGATCCAGAGCGCTCAATCATGTGGAAGCC 1017
 DB 718 LSNCSRLKMLILKNNDRSLTKFLQDAFQRLYLDLSNNKIQMIOKTSPEVNANLKM 777
 QY 1018 GCGCAGTGCAGCAAGCTGTTCCATCGCGTTTACACGCTGGCGCTTGAAG 1077
 DB 778 LLLHNRFLCTGDAVFWFVMMVNHTEVTIPLYATDVTCVGRGAKGQSVISLDLYTCBLD 837
 QY 1078 GACGTTTTCATAGCAGATATGTTGCTGCGCTGCGCTGCTTACTACAGCGCTTG 1137
 DB 838 TNLILPFLSLISVSLFMMVMTASHLYFWDVWYIYHFKAKIKYQRLISPDCCYDAFIV 897
 QY 1138 AATCAGACACATTAGTGGGCTCATGACTCTGCGAAATTTGGCTGAACGTTTGGG 1197
 DB 898 DTQDPATTEWVLAELVAKLEDPREKFNLCLEERDMLPGQVLENLQSISQISKTYFVM 957
 QY 1198 AGGAGTCCTTCTTCAAGGATAGAGAGATCTTCTGGGCAACAT 1245
 DB 958 TDVKAKTENFKIAFYLSHORLMDKVDVILILEKPFQSKFLQLRK 1005

RESULT 13
 US-10-141-759-358
 ; Sequence 358, Application US/10141759
 ; Publication No. US20030207361A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerlicsen, Mary E.

; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3330R1C197
 ; CURRENT APPLICATION NUMBER: US/10/141,759
 ; PRIOR APPLICATION: 2002-05-08
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 358
 ; LENGTH: 1049
 ; TYPE: PR1
 ; ORGANISM: Homo Sapien
 US-10-141-759-358

Query Match 1.8%; Score 40.8; DB 12; Length 1049;
 Best Local Similarity 5.4%; Pred. No. 0.028;
 Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAAGGCGCCGTCATTGACATGTAAGTGTGCGAAACCTGATGCTGAAG 597
 DB 298 HNSLSLGHVPRWRKINKLOEIDLSNPLAKEIDAKFLHFLPSLJQLDLSFNELOVR 357
 QY 598 GGTTCGGAGGTCCTTCCCAACATGCACTGCTTACAGATGAGACCGTTTGTGAC 657
 DB 358 ASNNLSQAFSSLKSLKILIRGVFKELKSFNLSPLHNLONLEVDLGTNFIKIANLSMF 417
 QY 658 CAATGTTCCATAGTGAAGCATGCGCTTCTTCCGAGAGATGTTCTTAAGCGGAAGTG 717
 DB 418 KQFRILKVIDLSVNKISPSGDSSEVGFCSNARTSVSEYEPQVLEQLHYFRYDKYASCRF 477
 QY 718 GCGGTCAATGCGCGTCATTTCCGAGCTCGTGGTGCAAGCAAGCTTTCATGCTGGG 777
 DB 478 KKEASMSVNESEYKQGLDLSKNSIFVKSDDPHLSFKLNSGNLISQTLNGSE 537
 QY 778 GTAGACAGTTCATATATATGAAGCAAGTATGATGCTGTTGAGCAAGCTTGTGCAT 837
 DB 538 FQPLAELRYLDFSNRDLHSTAFELHKLVDLSSNSHYQSEGITMMLNFTKLV 597
 QY 838 GCTTTCAGGAGCGCTCTAGTGTGCGCGAAATGGGGCGAGTTCCTCCGCT 897
 DB 598 LQKLMMNDNDISSSTRMESESLRLEFRGNHLDVLMREGDNRYQLDFNLLKLELDI 657
 QY 898 GCATTCGCTGTTTCTTCTCGAGCGCTTACGCGCTGCTTCGATGAGCGCGTCCA 957
 DB 658 SKNSLSLPSGVDPMPNLKNSLAKNGLSKSPKLOCLKYLETDLDSHNLTVPER 717
 QY 958 AATCCCGGACAGTGCAGACTTACTGTTGATCCAGAGCGCTCAATCATGTGGAAGCC 1017
 DB 718 LSNCSRLKMLILKNNDRSLTKFLQDAFQRLYLDLSNNKIQMIOKTSPEVNANLKM 777
 QY 1018 GCGCAGTGCAGCAAGCTGTTCCATCGCGTTTACACGCTGGCGCTTGAAG 1077
 DB 778 LLLHNRFLCTGDAVFWFVMMVNHTEVTIPLYATDVTCVGRGAKGQSVISLDLYTCBLD 837
 QY 1078 GACGTTTTCATAGCAGATATGTTGCTGCGCTGCGCTGCTTACTACAGCGCTTG 1137
 DB 838 TNLILPFLSLISVSLFMMVMTASHLYFWDVWYIYHFKAKIKYQRLISPDCCYDAFIV 897
 QY 1138 AATCAGACACATTAGTGGGCTCATGACTCTGCGAAATTTGGCTGAACGTTTGGG 1197
 DB 898 DTQDPATTEWVLAELVAKLEDPREKFNLCLEERDMLPGQVLENLQSISQISKTYFVM 957
 QY 1198 AGGAGTCCTTCTTCAAGGATAGAGAGATCTTCTGGGCAACAT 1245

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 20:32:25 ; Search time 132.658 Seconds
(without alignments)
7546.168 Million cell updates/sec

Title: US-09-434-837-10

Perfect score: 2268
Sequence: 1 atgcacgcttcaccctcctcct.....actaccgagtagaataag 2268

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing: first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	2217.8	97.8	24595	6	5428147-1	Patent No. 5428147
2	2162.8	95.4	2211	3	US-09-073-587-2	Sequence 2, Appl1
3	136.4	6.0	2279	1	US-08-008-216-8	Sequence 8, Appl1
4	136.4	6.0	2279	1	US-08-459-569-8	Sequence 8, Appl1
5	136.4	6.0	2279	1	US-08-458-831-8	Sequence 8, Appl1
6	136.4	6.0	21126	1	US-08-008-216-19	Sequence 19, Appl1
7	136.4	6.0	21126	1	US-08-459-569-19	Sequence 19, Appl1
8	136.4	6.0	21126	1	US-08-458-831-19	Sequence 19, Appl1
9	38.6	1.7	1575	3	US-09-230-188-2	Sequence 2, Appl1
10	37	1.6	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
11	37	1.6	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
12	36	1.6	1642	2	US-08-665-037-1	Sequence 1, Appl1
13	36	1.6	1642	2	US-08-666-067-1	Sequence 1, Appl1
14	36	1.6	1642	2	US-08-732-870-1	Sequence 1, Appl1
15	35.2	1.6	1032	4	US-09-415-277C-4	Sequence 4, Appl1
16	35.2	1.6	1032	4	US-09-415-277C-6	Sequence 6, Appl1
17	35.2	1.6	3593	4	US-09-404-627-3	Sequence 3, Appl1
18	35.2	1.6	4205	4	US-09-404-627-1	Sequence 1, Appl1
19	35	1.5	7218	4	US-08-232-463-14	Sequence 14, Appl1
20	34.6	1.5	1056	4	US-09-252-991A-8485	Sequence 8485, Ap
21	34.6	1.5	1830	4	US-09-252-991A-8526	Sequence 8526, Ap
22	34.6	1.5	1830	4	US-09-252-991A-8322	Sequence 8322, Ap
23	33.6	1.5	1617	4	US-09-252-991A-6497	Sequence 6497, Ap
24	33.6	1.5	1665	4	US-09-252-991A-6535	Sequence 6535, Ap
25	33.4	1.5	419	1	US-08-486-013-59	Sequence 59, Appl
26	33.4	1.5	419	2	US-08-482-279-59	Sequence 59, Appl
27	33.4	1.5	419	2	US-08-342-268-59	Sequence 59, Appl

28	33.4	1.5	419	3	US-09-015-968-59	Sequence 59, Appl
29	33.4	1.5	419	4	US-09-397-386-59	Sequence 59, Appl
30	33.4	1.5	1764	2	US-08-642-684-1	Sequence 1, Appl1
31	33.4	1.5	1764	3	US-09-157-397-1	Sequence 1, Appl1
32	33.4	1.5	1764	4	US-09-517-871-21	Sequence 21, Appl1
33	33.4	1.5	1764	4	US-08-642-684-3	Sequence 3, Appl1
34	33.4	1.5	1770	2	US-08-642-684-16	Sequence 16, Appl1
35	33.4	1.5	1779	1	US-08-208-036-15	Sequence 15, Appl1
36	33.4	1.5	1779	1	US-08-428-823-15	Sequence 15, Appl1
37	33.4	1.5	2631	1	US-08-208-036-13	Sequence 13, Appl1
38	33.4	1.5	2631	1	US-08-428-823-13	Sequence 13, Appl1
39	33.4	1.5	2631	1	US-08-510-215A-1	Sequence 1, Appl1
40	33.4	1.5	2631	4	US-09-517-871-19	Sequence 19, Appl1
41	33.4	1.5	3252	1	US-08-428-823-11	Sequence 11, Appl1
42	33.4	1.5	3252	1	US-08-428-823-11	Sequence 11, Appl1
43	33.4	1.5	3317	2	US-08-633-476-1	Sequence 11, Appl1
44	33.4	1.5	7218	1	US-08-232-463-14	Sequence 14, Appl1
45	33.4	1.5	536165	4	US-09-214-808-1	Sequence 1, Appl1

ALIGNMENTS

```
RESULT 1
5428147-1
; Patent No. 5428147
; APPLICANT: BARKER, RICHARD F.; KEMP, JOHN D.
; TITLE OF INVENTION: OCTOPINE T-DNA PROMOTERS
; NUMBER OF SEQUENCES: 17
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/91,538
; FILING DATE: 13-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 869,216
; FILING DATE: 13-APR-1992
; APPLICATION NUMBER: 869,216
; FILING DATE: 13-APR-1992
; APPLICATION NUMBER: 440,432
; FILING DATE: 21-NOV-1989
; APPLICATION NUMBER: 553,786
; FILING DATE: 19-NOV-1983
; APPLICATION NUMBER: 741,034
; FILING DATE: 06-AUG-1991
; APPLICATION NUMBER: 144,775
; FILING DATE: 20-JAN-1988
; APPLICATION NUMBER: 485,614
; FILING DATE: 15-APR-1983
; APPLICATION NUMBER: 713,624
; FILING DATE: 10-JUN-1991
; APPLICATION NUMBER: 260,574
; FILING DATE: 21-OCT-1988
; APPLICATION NUMBER: 848,733
; FILING DATE: 01-APR-1986
; APPLICATION NUMBER: 535,354
; FILING DATE: 26-SEP-1983
; SEQ ID NO: 1
; LENGTH: 24595
5428147-1
Query Match 97.8%; Score 2217.8; DB 6; Length 24595;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2250; Conservative 0; Mismatches 17; Indels 2; Gaps 2;
```

OY 1 ATGTGAGCTTCACTCTCTGATTAACGATGATCATCTCCCAACAAATGTGTGAT 60
|||
Db 5809 ATGTGAGCTTCACTCTCTGATTAACGATGATCATCTCCCAACAAATGTGTGAT 5868
|||
OY 61 CTGACATGTTGATTAAGCGGATGATTAAGCGGATGATTAAGCGGATGATTAAGCGGAT 120
|||
Db 5869 CTGACATGTTGATTAAGCGGATGATTAAGCGGATGATTAAGCGGATGATTAAGCGGAT 5928
|||
OY 121 CGAAGAGCTTCAAGGAGGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTA 180
|||

Dh 5929 CGAAGACCTTCTAGGGAAGAGATTACTCAATCTCCACGAGTGCAGCGCTGGGTTA 5988
Qy 181 GCTTGCAGAAAGCTGGCCGATGGTCCGTTCCCGAGATCTCAGCTGGTGGAAAGTAGCA 240
Db 5989 GCTTGCAGAAAGCTGGCCGATGGTCCGTTCCCGAGATCTCAGCTGGTGGAAAGTAGCA 6048
Qy 241 GTTCTCTCGCTTATATCTATTTATTTGGCAAAATTTGGGCGGATCTTGATCGAAA 300
Db 6049 GTTCTCTCGCTTATATCTATTTATTTGGCAAAATTTGGGCGGATCTTGATCGAAA 6108
Qy 301 CTTTGGCGGGGCAACAGTAGTGTCTGTTGCCATGCACTTGGACACCATTTGCATG 360
Db 6109 CTTTGGCGGGGCAACAGTAGTGTCTGTTGCCATGCACTTGGACACCATTTTGCATG 6168
Qy 361 GATTTTCCGCAACCACTAATCCAGCCCTGTTTGGTGAACCGTTAAACATGCGA 420
Db 6169 GATTTTCCGCAACCACTAATCCAGCCCTGTTTGGTGAACCGTTAAACATGCGA 6228
Qy 421 CCGATTGATCTTAGTCACTTTCGTGGCCATTTCAATCTTAAGACTGGCGCTTTGCAAC 480
Db 6229 CCGATTGATCTTAGTCACTTTCGTGGCCATTTCAATCTTAAGACTGGCGCTTTGCAAC 6288
Qy 481 CTGCAATGCGCGCTGTACGAGAAATGCGACGATGAATGCGTTACCGGTTTACCATACC 540
Db 6289 CTGCAATGCGCGCTGTACGAGAAATGCGACGATGAATGCGTTACCGGTTTACCATACC 6348
Qy 541 CTTGAAGAGGCGCGTCCATTTGACATGATAGCTTATATGTTGAAAACCTGATGCTGAAAGGT 600
Db 6349 CTTGAAGAGGCGCGTCCATTTGACATGATAGCTTATATGTTGAAAACCTGATGCTGAAAGGT 6408
Qy 601 TCGGCAAGGTTCTTTCCACCAATCGACTGCTCTACAGACTGCAAGCCGTTTGTGACCA 660
Db 6409 TCGGCAAGGTTCTTTCCACCAATCGACTGCTCTACAGACTGCAAGCCGTTTGTGACCA 6468
Qy 661 TGTTCGATATGAGACGAGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAATGGCG 720
Db 6469 TGTTCGATATGAGACGAGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAATGGCG 6528
Qy 721 GTATATGGCCCTGGCATTTCCGACTGCTGATGGTGGCAACGAACTGCTTATGCTGGGTTA 780
Db 6529 GTATATGGCCCTGGCATTTCCGACTGCTGATGGTGGCAACGAACTGCTTATGCTGGGTTA 6588
Qy 781 GACGATGTCAATATATGAGCAAGATGATGATGTTGGAGCAAGCTTTGGTCAATGCT 840
Db 6589 GACGATGTCAATATATGAGCAAGATGATGATGTTGGAGCAAGCTTTGGTCAATGCT 6648
Qy 841 TTCAGGACGCTCTAGTGTGTGTGGCCGAATGGGGCGATGCGGATTTCTCTGCTGCA 900
Db 6649 TTCAGGACGCTCTAGTGTGTGTGGCCGAATGGGGCGATGCGGATTTCTCTGCTGCA 6708
Qy 901 TTTCTGCTTTTCTTCTCTGAGCCGTTACGCGCTGTCTTGATGAGCCGTTCCCAAT 960
Db 6709 TTTCTGCTTTTCTTCTCTGAGCCGTTACGCGCTGTCTTGATGAGCCGTTCCCAAT 6768
Qy 961 CCGGCGCAGTGCACCTTACTTGGTCTACCAAGCGCTCCCAATCATGTGAAAACCGGG 1020
Db 6769 CCGGCGCAGTGCACCTTACTTGGTCTACCAAGCGCTCCCAATCATGTGAAAACCGGG 6828
Qy 1021 CAGCTGCGACGAAAGCTTTCATCGCGTTTACACGCGTTGGCGTGGCTTCTTGAAGAC 1080
Db 6829 CAGCTGCGACGAAAGCTTTCATCGCGTTTACACGCGTTGGCGTGGCTTCTTGAAGAC 6888
Qy 1081 GGTTTTCAAGAGAGATATTTGTGTGGCTTCCGCTGTCTTATCTCAGGCTTTGAAA 1140
Db 6889 GGTTTTCAAGAGAGATATTTGTGTGGCTTCCGCTGTCTTATCTCAGGCTTTGAAA 6948
Qy 1141 TCAGGACATTTAGTGGGCTCATGACTCCGCGCAATTTGGCTGAACGTTTGGGAGG 1200
Db 6949 TCAGGACATTTAGTGGGCTCATGACTCCGCGCAATTTGGCTGAACGTTTGGGAGG 7008
Qy 1201 GAGTCTTCTTCTCAGGAGATAGAGAGATCTTCTTGGGCAACATCTCTCTGTGTGT-AA 1259
Db 7009 GAGTCTTCTTCTCAGGAGATAGAGAGATCTTCTTGGGCAACATCTCTCTGTGTGTGAA 7068

Qy 1260 ACATGAGTTTCTCATGATTTGGGACCTATTCAAGCTAATGGGAATAGGATCTGGCGG 1319
Db 7069 ACATGAGTTTCTCATGATTTGGGACCTATTCAAGCTAATGGGAATAGGATCTGGG-GG 7127
Qy 1320 GTTGTGCTCAGTTTGTGAAAGCGGGTTTATTTAGATCTTCCGCTGGTGTATCAACGGATA 1379
Db 7128 GTTGTGCTCAGTTTGTGAAAGCGGGTTTATTTAGATCTTCCGCTGGTGTATCAACGGATA 7287
Qy 1380 TGAAGAAATCAGCGGATGTCCCTGAAGAAATCTCAGAACTTCCAGCTGCGATCCGATC 1439
Db 7188 TGAAGAAATCAGCGGATGTCCCTGAAGAAATCTCAGAACTTCCAGCTGCGATCCGATC 7247
Qy 1440 TGAAGTTTAAACGAGTGTCTGTGAGCCAGGCAATATGCAATGCTCAAGTCAAGGCGAT 1499
Db 7248 TGAAGTTTAAACGAGTGTCTGTGAGCCAGGCAATATGCAATGCTCAAGTCAAGGCGAT 7307
Qy 1500 TCAGAGGAAAGACAAATAATTAAGATTAAGGCTTAAGAGCGGGATATCTGAATTTATGA 1559
Db 7308 TCAGAGGAAAGACAAATAATTAAGATTAAGGCTTAAGAGCGGGATATCTGAATTTATGA 7367
Qy 1560 TAAAGTGTGTCACTCTGACTCGCAATATCCAACTCAGGCTATGCTGCAATGCGA 1619
Db 7368 TAAAGTGTGTCACTCTGACTCGCAATATCCAACTCAGGCTATGCTGCAATGCGA 7427
Qy 1620 TACCAATTTTTTCAAGCAACAGTGAACCAAGGCTTGAATTAACAGCCATATGACAGATC 1679
Db 7428 TACCAATTTTTTCAAGCAACAGTGAACCAAGGCTTGAATTAACAGCCATATGACAGATC 7487
Qy 1680 GTCAAAATCTTCTCTGATGACTGAAGCAAAATTTCTGTTATAGACATATCTCCGCTTGG 1739
Db 7488 GTCAAAATCTTCTCTGATGACTGAAGCAAAATTTCTGTTATAGACATATCTCCGCTTGG 7547
Qy 1740 TGTCTCATGAGAGGATGCGAAAGCAGTGTATGCTTGAATATGATGATGCGAGATCC 1799
Db 7548 TGTCTCATGAGAGGATGCGAAAGCAGTGTATGCTTGAATATGATGATGCGAGATCC 7607
Qy 1800 GAATGTAAAGGCTAGTGTCTATCATATATCATGAGAGAGAGATCCCAAGCTGTT 1859
Db 7608 GAATGTAAAGGCTAGTGTCTATCATATATCATGAGAGAGAGATCCCAAGCTGTT 7667
Qy 1860 GGGGCTCCCCGACAAAGAAAGAGCAATATGCTGTGCGGAGCGCAATTTGAGATCTTT 1919
Db 7668 GGGGCTCCCCGACAAAGAAAGAGCAATATGCTGTGCGGAGCGCAATTTGAGATCTTT 7727
Qy 1920 CCGGCGCTTGGCCAGACCTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATTC 1979
Db 7728 CCGGCGCTTGGCCAGACCTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATTC 7787
Qy 1980 ACATGATGAGCTTACAGAGAGATGCGGGGAGCTTCAAACTCAACCGGCGTGGTGA 2039
Db 7788 ACATGATGAGCTTACAGAGAGATGCGGGGAGCTTCAAACTCAACCGGCGTGGTGA 7847
Qy 2040 GGAATTTTATTTGAAGAACTTTCTTCAAGCACTGACACGCGTAAATGATACCGAGT 2099
Db 7848 GGAATTTTATTTGAAGAACTTTCTTCAAGCACTGACACGCGTAAATGATACCGAGT 7907
Qy 2100 TTAATTTGGGCTTCAAGTGTCTTCAAGTGTGATGAGTGTGAGGCTGCTATTCAGAC 2159
Db 7908 TTAATTTGGGCTTCAAGTGTCTTCAAGTGTGATGAGTGTGAGGCTGCTATTCAGAC 7967
Qy 2160 CCGGCTTAAACGCGCTGCTGCAATTAATCAATTTGAGAGGCAATTTGGCAAGGCA 2219
Db 7968 CCGGCTTAAACGCGCTGCTGCAATTAATCAATTTGAGAGGCAATTTGGCAAGGCA 8027
Qy 2220 TCTCTGAAACACTTTGAGAGATATACTACCCAGTAGTAATTAG 2268
Db 8028 TCTCTGAAACACTTTGAGAGATATACTACCCAGTAGTAATTAG 8076

RESULT 2
US-09-073-587-2
; Sequence 2, Application US/09073587

Patent No. 626552
GENERAL INFORMATION:
APPLICANT: Li, Yi
TITLE OF INVENTION: Transgenic Seedless Plants
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Wimer and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,587
FILING DATE: 06-MAY-1998
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/045,725
FILING DATE: 06-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 4-97
TELEPHONE: (303) 499-8089
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2211 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 4..2205
US-09-073-587-2

Query Match 95.4%; Score 2162.8; DB 3; Length 2211;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

QY 64 ACAATGTCGATTAAGCGGATGGAATGGACGAGGTTCCGATGCTTCTTAGAACCA 123
DB 1 ACATGATCAATTAAGCGGATGGAATGGACGAGGTTCCGATGCTTCTTAGAACCA 60
QY 124 GAAGCTTCTAGGGAAGAGATTACTCAAGCTCCACCGAGTGCAGCGCTGGGTTAGCT 183
DB 61 GAAGCTTCTAGGGAAGAGATTACTCAAGCTCCACCGAGTGCAGCGCTGGGTTAGCT 120
QY 184 TGCAGAAAGCTGCGCATGTGCTTCCCGAGATCTCAGCTGTGGAAGGTAGCAATT 243
DB 121 TGCAGAAAGCTGCGCATGTGCTTCCCGAGATCTCAGCTGTGGAAGGTAGCAATT 180
QY 244 CTCTCGGCTTATCTATATTGGCAAGAAATCTGGGGGGGATTAATTGAATGAACCT 303
DB 181 CTCTCGGCTTATCTATATTGGCAAGAAATCTGGGGGGGATTAATTGAATGAACCT 240
QY 304 TGGGCGCGGCAACAGTAGTGTCTCTTGCATGCACTTGACACCAATTTGCATGAT 363
DB 241 TGGGCGCGGCAACAGTAGTGTCTCTTGCATGCACTTGACACCAATTTGCATGAT 300
QY 364 TTCTCCGAAGCACTAATCCAAAGCCGTGTTTGTGTAAGCGGTAAAGATGTGACCG 423
DB 301 TTCTCCGAAGCACTAATCCAAAGCCGTGTTTGTGTAAGCGGTAAAGATGTGACCG 360

QY 424 ATTGATCTTAGTATTCGTGCGCATTTCAATCTTAAGACTCCGGCTTTGCAACCTG 483
DB 361 ATTGATCTTAGTATTCGTGCGCATTTCAATCTTAAGACTCCGGCTTTGCAACCTG 420
QY 484 CCAATCCCGTGTAGAGAAATGACAGATGAATGCGGTTTACCAATACCTT 543
DB 421 CCAATCCCGTGTAGAGAAATGACAGATGAATGCGGTTTACCAATACCTT 480
QY 544 GAAGGGGCGGTGCATTTGACATGATAGCTTATGCTGCAAACTGATTCGGAAGGTTG 603
DB 481 GAAGGGGCGGTGCATTTGACATGATAGCTTATGCTGCAAACTGATTCGGAAGGTTG 540
QY 604 GCAGTTCTCTTCCAAATCGACTTCTCTACGACTGAGACGCTTTTTCACCATGT 663
DB 541 GCAGTTCTCTTCCAAATCGACTTCTCTACGACTGAGACGCTTTTTCACCATGT 600
QY 664 TCCGATAGTGAAGGATGCGGCTTCTTCCGAGATGTTCTTAAAGCGAAGTGGCGTC 723
DB 601 TCCGATAGTGAAGGATGCGGCTTCTTCCGAGATGTTCTTAAAGCGAAGTGGCGTC 660
QY 724 ATTGGGCGTGCATTTCCGCACTGCTGTGTGCGCAAACTGCTTATGCTGGGTTAGC 783
DB 661 ATTGGGCGTGCATTTCCGCACTGCTGTGTGCGCAAACTGCTTATGCTGGGTTAGC 720
QY 784 GATGTACATATATGAGCAAGTATCGTTTGAGAGCAAGCTTGTGCATGCTTTC 843
DB 721 GATGTACATATATGAGCAAGTATCGTTTGAGAGCAAGCTTGTGCATGCTTTC 780
QY 844 AAGGACGCTCTAGTGTGCGGCAAAATGGGGCGATCGATTTCTCTGCTGATTC 903
DB 781 AAGGACGCTCTAGTGTGCGGCAAAATGGGGCGATCGATTTCTCTGCTGATTC 840
QY 904 TGCTTTTCTTCTCTGAGGCTTACGGCTGTCTTCTGATAGGCGCTTCCAAATCCC 963
DB 841 TGCTTTTCTTCTCTGAGGCTTACGGCTGTCTTCTGATAGGCGCTTCCAAATCCC 900
QY 964 GGCACATGACACTTACTGTGCTTACCAAGGCGTCAATATGAGAAAGCGGGCAG 1023
DB 901 GGCACATGACACTTACTGTGCTTACCAAGGCGTCAATATGAGAAAGCGGGCAG 960
QY 1024 CTGCGACGAGCTGTTCATGCGCTTTTCAACGCTGTGCGTCTTCTTGAAGACGGT 1083
DB 961 CTGCGACGAGCTGTTCATGCGCTTTTCAACGCTGTGCGTCTTCTTGAAGACGGT 1020
QY 1084 TTTCATGACGATATTTGTGCTTCTGCTGTCTTCTGATTAATCTAGGCTTGAATCA 1143
DB 1021 TTTCATGACGATATTTGTGCTTCTGCTGTCTTCTGATTAATCTAGGCTTGAATCA 1080
QY 1144 GGAACATTAAGTGGGCTGATGCTCTGCGCAAAATTTGGCTGAACCGTTTCGGGAGGAG 1203
DB 1081 GGAACATTAAGTGGGCTGATGCTCTGCGCAAAATTTGGCTGAACCGTTTCGGGAGGAG 1140
QY 1204 TCTTCTCTTCAAGGATAGAGAGATCTTCTGCGCACATCTCTCTGCTGTGTT -AACA 1262
DB 1141 TCTTCTCTTCAAGGATAGAGAGATCTTCTGCGCACATCTCTCTGCTGTGTT -AACA 1200
QY 1263 TGAAGTTTCTCATGATTTGGAACCTTCAAGCTTATGGAATAGAGATCTGCGGGGTT 1322
DB 1201 TGAAGTTTCTCATGATTTGGAACCTTCAAGCTTATGGAATAGAGATCTGCGG -GGGTT 1259
QY 1323 TGGTCAGTTTGAAGAGGGGTTTATGAGATCTCTGCGCTGTGCTATGAAGGATATGA 1382
DB 1260 TGGTCAGTTTGAAGAGGGGTTTATGAGATCTCTGCGCTGTGCTATGAAGGATATGA 1319
QY 1383 AGAAATCAGCGAGTGTGCTGGAAGAGATCTCAGAACTTCCAGCTGCGATGCACTGA 1442
DB 1320 AGAAATCAGCGAGTGTGCTGGAAGAGATCTCAGAACTTCCAGCTGCGATGCACTGA 1379
QY 1443 AGTGTAAAGGTTGTCTGTAGCGAGCATATGCAATGTTCAAGTCAAGGCGGATTTCA 1502
DB 1380 AGTGTAAAGGTTGTCTGTAGCGAGCATATGCAATGTTCAAGTCAAGGCGGATTTCA 1439
QY 1503 GAAAGAAAAGCAAAAATTAAGATTAAGGCTTAAAGCGGAGATATCTGAACCTTATGATGA 1562

Db	1440	GAAGGAAAGACAAATAATTAAGTAAAGCTTAAAGCGGAGTATCTGAACCTTTATGATTA	1499
Qy	1563	GGTGTGTGTCACATCTGGACTCGCAATATCCAACTCAGGCATTTGCTGACATGCGATAC	1622
Db	1500	GGTGTGTGTCACATCTGGACTCGCAATATCCAACTCAGGCATTTGCTGACATGCGATAC	1559
Qy	1623	CAATATTTTTCAGGACACAGTGAACCAAGGGGTTGATTAACAGGCATATGACAGATGCTC	1682
Db	1560	CAATATTTTTCAGGACACAGTGAACCAAGGGGTTGATTAACAGGCATATGACAGATGCTC	1619
Qy	1683	AAAACCTTCTCCGTATGACTGACCAAAATTTCTGGTTAGACCATATCCTCCGCTTGTGT	1742
Db	1620	AAAACCTTCTCCGTATGACTGACCAAAATTTCTGGTTAGACCATATCCTCCGCTTGTGT	1679
Qy	1743	CCTCATGAGCGGATCGCAAAAAGCATGTATTGCTTGGACTATGATGTCGACGATCCGAA	1802
Db	1680	CCTCATGAGCGGATCGCAAAAAGCATGTATTGCTTGGACTATGATGTCGACGATCCGAA	1739
Qy	1803	TGCTAAAGGCTATAGTGTCACTCACTTATTAACATGGGAGAGACATCCCAACAAGCTGTGCG	1862
Db	1740	TGCTAAAGGCTATAGTGTCACTCACTTATTAACATGGGAGAGACATCCCAACAAGCTGTGCG	1799
Qy	1863	GGTCCCGGACAAAAAAGAGCATATATGTCTGCTCGGGAACGCAATTTTCAGATCTTTCC	1922
Db	1800	GGTCCCGGACAAAAAAGAGCATATATGTCTGCTCGGGAACGCAATTTTCAGATCTTTCC	1859
Qy	1923	GGCGTTTTCAGGACCTATTTCTGCTGCTGCTGATTAACGACCAAAATGTTATTCACA	1982
Db	1860	GGCGTTTTCAGGACCTATTTCTGCTGCTGCTGATTAACGACCAAAATGTTATTCACA	1919
Qy	1983	TGATTTGCTTACACAGACGAAATGCGGGGGAGCTTTAAACTCAACCGGGGTGTGAGGA	2042
Db	1920	TGATTTGCTTACACAGACGAAATGCGGGGGAGCTTTAAACTCAACCGGGGTGTGAGGA	1979
Qy	2043	TTTTTATTTCTGAAGAACTTTTCTTTCAAGCACTGACACGCGCTATATACCGGAGTTTA	2102
Db	1980	TTTTTATTTCTGAAGAACTTTTCTTTCAAGCACTGACACGCGCTATATACCGGAGTTTA	2039
Qy	2103	CTTGGCGGGTTGACAGTTGTTCTTTACAGGTGATGGGTGAGGGGTCCTATTCCAGCCGC	2162
Db	2040	CTTGGCGGGTTGACAGTTGTTCTTTACAGGTGATGGGTGAGGGGTCCTATTCCAGCCGC	2099
Qy	2163	GTCTAAAGCGCCCTCTGTGCAATTATTCACAAATTTGGAGGCACTTTTGGCAAAAGGCAATCC	2222
Db	2100	GTCTAAAGCGCCCTCTGTGCAATTATTCACAAATTTGGAGGCACTTTTGGCAAAAGGCAATCC	2159
Qy	2223	TCTGGAACACACTCTTGAAGAGATATTAACCGCAGTAGAAATTAG 2268	
Db	2160	TCTGGAACACACTCTTGAAGAGATATTAACCGCAGTAGAAATTAG 2205	
RESULT 3			
US-08-008-216-8			
Sequence 8, Application US/08008216			
Patent No. 5368887			
GENERAL INFORMATION:			
APPLICANT: Slighton, Jerry L.			
APPLICANT: Tepfer, David A.			
TITLE OF INVENTION: R1 T-DNA Promoters			
NUMBER OF SEQUENCES: 19			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: GRAY, CARY, AMES & FRYE			
STREET: 401 B Street, Suite 1700			
CITY: San Diego			
STATE: California			
COUNTRY: USA			
ZIP: 92101-4297			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0 Version #1.25			

```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/08/008,216
3 FILING DATE: 25-JAN-1993
4 CLASSIFICATION: 800
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: US 06/725,368
7 FILING DATE: 22-APR-1985
8 ATTORNEY/AGENT INFORMATION:
9 NAME: Barthorst, Marjie W.
10 REGISTRATION NUMBER: 36,740
11 REFERENCE/DOCKET NUMBER: P1020U51
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: (619) 699-2700
14 TELEFAX: (619) 236-1048
15 INFORMATION FOR SEQ ID NO: 8:
16 SEQUENCE CHARACTERISTICS:
17 LENGTH: 2279 base pairs
18 TYPE: nucleic acid
19 STRANDEDNESS: double
20 TOPOLOGY: linear
21 MOLECULE TYPE: DNA (genomic)
22 ORIGINAL SOURCE:
23 ORGANISM: Agrobacterium rhizogenes
24 STRAIN: Strain A4
25 IMMEDIATE SOURCE:
26 LIBRARY: Convolvulus arvensis plant cells
27 CLONE: Clone 7
28 FEATURE:
29 NAME/KEY: misc feature
30 LOCATION: 1..2279
31 OTHER INFORMATION: /label= ORF8
32 OTHER INFORMATION: /note= "Sequence ORF8 corresponds to bases 6609
33 OTHER INFORMATION: through 8888 of Seq. ID No. 5366887 19. "
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970

```

Query Match	5.0%	Score 136.4	DB 1	Length 2219
Similarity	45.9%	Pred. No. 8.7e-37		
Best Local	0	Mismatches 721	Indels 14	Gaps 4
Match 623: Conservative				
QY	623	TCGACGTGCTCTACGACCTCGACGACCGTTTTTTTGAACCAATGTTCCGATGTTAGTAGAGCATGCG	682	
DB	665	TTGACCTCGCGGTACGACTATACAGGTTTTTCTCCAGAGTGGGATGCCATATATGTGTGCG	724	
QY	683	GCTTCTTTCCGGAGGATGTTCTTAAGCCGAAAGTGGCGGTCATTGGCGCTGGCAATTTCCG	742	
DB	725	GGCATTTATCCGGAAGGGGACCACTACTACTAGTGTTCGCATCGTAGAGGCGGGCTGTCTG	784	
QY	743	GACTCGTGTGTGGCAAGCAATCGTTCTTACGTGTGGGGTAGACAGTGTTCAAATATATGAG	802	
DB	785	GCTTGTGTGCTGCACAGAACTACTTGGCGCTGGCGTCAAGAAATACCTTTTTCGATA	844	
QY	803	CAACTGATTCGTTGTGAGGCA---GCTTTGGTCAATGCTTTACGGGAGCGTCTACTG	859	
DB	845	CCGTTGATGAGATCCGATGTTTGGGGCATGCCGATGCCAAAGCGGACGCTTACCGAG	904	
QY	860	TCGTTGCCGAATGGGGCGATGCGATTTCCTCCGTGTCATTCGTGTTTTTCTTCC	919	
DB	905	CCTTAGCGTCGTTGGGTGTATGCTTCTTCGCCAACCAACTTTGGCTGTATCTATC	964	
QY	920	TCGAGCGTTACGCGCTGTCTTGATGAGGCCGTTTCCAAATCCGGGACAGTCCGACTT	979	
DB	965	TGGATTAAGTTAAGATTCCGTCCAGCCCTTGTTTTCTTGCCGGCAACACACACAG	1024	
QY	980	ACTTGTGTACCAAGGCGTCCAAATACATGTGGAAGCCGGGACGTGCAACCGAAGCTGT	1034	
DB	1025	CACATAATTTCCGCGAGAAAGCATACGATGCGACGCGGGCAAGCTTCGCGGGGATAT	1084	
QY	1040	TCCATCGGTTTACCAACGCTTGGGGGTGCTTCTTGAAGAGGTTTTTCATGAGCAGATA	1099	
DB	1085	TTACGCGGGTACATGTGCGATGAGAACACTACTCTTACCAAGGGGTGTAAACGAAATGCA	1144	
QY	1100	TTTGTTGGGCTTGGCGTGTGCTTATCTACAGGCTTGAATTCAGGACACACTTAAGTGGG	1155	

Db	1145	GGAGCTGATGGCTCCGATNGATATCTTTTCATGTTAAAGAGCGTGTGTATGAG	1204
Qy	1160	CTCATGATCTCTGGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTTTCAGGGA	1219
Db	1205	CCTCAGAAAGCACGCGACGGCTTTGGCTCCGAGAGTTGGGAAAAATTCACCTTTCATGCCGTTT	1264
Qy	1220	TAGAGAGATCTTCTGTGGGACACATCTCTCCGTGGGT-----AAACATGAGATTTCCTC	1275
Db	1265	TGCTCGAATCTTCACTGCTGTGTATTTGAGTCCGTGGGCGAAGGCAAGGCAACACCCC	1324
Qy	1276	ATGATTTGGGACCTATTCAGCTAAATGGGAATAGATCTGCGGGGTTTGTCTCAGTTT	1335
Db	1325	ATGATTTGAGGCTTTCCGGGATACTGAGTTGGGATACGCCGAGTTTCGTCC-TATTAC	1383
Qy	1336	GAAGCGGCTTATTTAGATCTCTCCGCTTGGTCATCAACGGAATGAGAAATACGCGG	1395
Db	1384	AACGTTGTGTTTTCACAGATCCTGGACTGATTATCAATGGCTACGAGGAGACAGCAT	1443
Qy	1396	ATGTGCTTGAAGAAATCTCAGAACTTCCACGTCGATTCGATCTGAAGTGTTAACGGT	1455
Db	1444	CTTTCATATTTGTTGGGTTTCAACTTTTGGACGCTGATGCGCATTTGAAATTTTCCAGAA	1503
Qy	1456	GTCGTCTGAGCGACGCGATATGCGATATGCAATGTCAGGAGGCGATTACAGAGAAAGACA	1515
Db	1504	AGCATTGCGAAAGCACGACTCTGTTTATCTCCGTGCGTGAATAGCCAGAGAGGCGGG	1563
Qy	1516	AAAAATAAAGATTAAGCTTTAAGACGCGGATATCTGAACCTTATATGATTAAGTGGTGTCA	1575
Db	1564	AGATTGAAGATATGTTGTAACACGGTCATTGCGGTGTTTTTGAACAGTCAATATTGGC	1623
Qy	1576	TCTGGACTCGCAATATATCCAACTCAGGCAATTGCTGACATGCGATACCAATATTTTCAG	1635
Db	1624	GGCAGTGTGAGGCGCGCTACAGTTGATTAACAGACTGGCCGGGGATGAGACTTCTTCAGC	1683
Qy	1636	GCACAGTGAACCAAGCGGTTGATTAACAGCAATAGACAGATGTCTAAACTCTCTCTG	1695
Db	1684	TACATATTCGAACCCGCGGTGGAACCTGCTGCGCTGTCAATTCCAGCACTCTTATG	1743
Qy	1696	ATGACTGAACAAAAATTCGTGTAGACCATATCTCCCGCTTGTGTCCTCATGGACGG	1755
Db	1744	GTCAGAGAGCAAAAGTTTGTGGTTAATCTCCGGCATCCAGCAGATATAGAACCGATGG	1803
Qy	1756	ATCGCAAAAGCAGTGTATTTGCTCGGACTATGATGATCGAGATCCGAATGTAAAGTCTA	1815
Db	1804	CTTGTCCGTGAGCTGTGTGATTCATCGAATG-----CCAGCTGAGAGAGGCTT	1857
Qy	1816	GTCGTCAATCAGTATATACATGAGAGACACTCCCAAGCTGTGGCGGTCCCGACAA	1875
Db	1858	GTCGTTTTCATATGCTTTGTGATGATCATATCTCCGCGATCGACATCATGACAG	1917
Qy	1876	AAAGAGCATATATGCTCTGTCGCGGACCGCAATTTGAGATCTTTCGCGGCTTTGCCAG	1935
Db	1918	AAAGGACGCTCTTGGAATTTGGTCAAGGAGCTGTGCTGCTTTCCTGAACCTGGCTTG	1977
Qy	1936	CACCTAATTCCTGCTGCTGCTGATTCGACCAAAATGT	1973
Db	1978	CACCTGTCCAGTCAAGAGACTACGAACGATATGT	2015
RESULT 4			
US-08-459-569-8			
Sequence 8, Application US/08459569			
Patent No. 5543501			
GENERAL INFORMATION:			
APPLICANT: Slightom, Jerry L.			
APPLICANT: Tepler, David A.			
TITLE OF INVENTION: R1 T-DNA Promoters			
NUMBER OF SEQUENCES: 19			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: GRAY, CAR, AMES & FRYE			
STREET: 401 B Street, Suite 1700			
CITY: San Diego			
STATE: California			

```

1 COUNTRY: USA
2 ZIP: 92101-4297
3 COMPUTER READABLE FORM:
4 MEDIUM TYPE: Floppy disk
5 COMPUTER: IBM PC compatible
6 OPERATING SYSTEM: PC-DOS/MS-DOS
7 SOFTWARE: Patent Release #1.0, Version #1.25
8 CURRENT APPLICATION DATA:
9 APPLICATION NUMBER: US/08/459,569
10 FILING DATE: 02-JUN-1995
11 CLASSIFICATION: 530
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: US 08/008,216
14 FILING DATE: 25-JAN-1993
15 APPLICATION NUMBER: US 06/725,368
16 FILING DATE: 22-APR-1985
17 ATTORNEY/AGENT INFORMATION:
18 NAME: Barnhorst, Marnie W.
19 REGISTRATION NUMBER: 36,740
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: (619) 699-2700
22 TELEFAX: (619) 236-1048
23 INFORMATION FOR SEQ ID NO: 8:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 2279 base pairs
26 TYPE: nucleic acid
27 STRANDEDNESS: double
28 TOPOLOGY: linear
29 MOLECULE TYPE: DNA (genomic)
30 ORIGINAL SOURCE:
31 ORGANISM: Agrobacterium rhizogenes
32 STRAIN: Strain A4
33 IMMEDIATE SOURCE:
34 CLONE: Clone 7
35 LIBRARY: Convolvulus arvensis plant cells
36 FEATURE:
37 NAME/KEY: misc feature
38 LOCATION: 1..2279
39 OTHER INFORMATION: /label= ORF8
40 OTHER INFORMATION: /note="Sequence ORF8 corresponds to bases 6609
41 Patent No. 5543501
42 OTHER INFORMATION: through 8888 of Seq. ID No. 5543501 19. "
43 US-08-459-569-8
44
45 Query Match 6.0%; Score 136.4; DB 1; Length 2279;
46 Best Local Similarity 45.9%; Pred. No. 8,7e-37;
47 Matches 623; Conservative 0; Mismatches 721; Indels 14; Gaps 4;
48
49 QY 623 TCACATGCTCTACGACTGCAGACCGTTTTTGTGACCAATGTTCCGATAGTGGAGGATCG 682
50 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 Db 665 TTGACCGCGGATGACCTATTCAGGTTTTTCTCCAGAGTGGATGCCCATATGGTGTGG 724
52 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
53 QY 683 GCTTCTTTCCGAGAGATGTTCTTAAAGCCGAAGTGGCGTTCATTGGCGCTGGCATTTCCG 742
54 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
55 Db 725 GGCATTATCCGGAAGGCGACACTACTACTAGTGTGCGCATGTAGAGGCGGGCTGTG 784
56 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
57 QY 743 GACTCGTGTGGCCAAAGCAACTGCTTCAATGCTGGGGGTAGACAGATTTCATATATAGAG 802
58 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
59 Db 785 GCCTTGTTGTCTGCACAGAACTACTTTCGCGCTGCCTCAAGGAATACATCTTTTTCGATA 844
60 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 QY 803 CAAGTATCGTGTGAGGCA---GCTTGGTGCATATGCTTTCAGGAGCGTCTCTAGT 859
62 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63 Db 845 CCGTTATGAGATCCGATGTTTGGGGCATGCCGATGCCAAACGGCGACTACACAG 904
64 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
65 QY 860 TCGTGCCGAATGGGGGCGATGCGATTTCCTCGTGCATTCGTGTTTCTTCTCC 919
66 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 Db 905 CATTGACGTGCTTCGGGTGTCATGCTTTTCCTCCGCAACCAACTTTGCTGTCATATATC 964
68 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
69 QY 920 TCGAGCGTTACGCGCTGTCTTCGATAGAGCCGTTCCCAATCCCGGCAACAGTCAACTT 979
70 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
71 Db 965 TGAATAGATTAGAAATCCGTCACGCTTGTTTCTTGTGCGGCAACAGCACACAG 1024
72 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 980 ACTTGCTACCAAGGCGCTCAATACATGTGAAAGCCGGGAGCTGCCACCGAAGCTGT 1039
DB 1025 CACTATATTTCCGCGACAAACGCTACGATGCGACGGGGGCAAGCTCCCGGGGATAT 1084
QY 1040 TCCATGCGCTTTACAAAGGTTGGCGCTGCTTTGAAGGACGGTTTTCATAGCGAGATA 1099
DB 1085 TTGAGCGGGTACATGTGCGATGGAAGACATCTACTCTACAAAGGTGTGAACGGAAATGCA 1144
QY 1100 TTGTTGGCTTGGCGCTGTGCTATTTACTCAGGCGCTTGAATCAGACACATTAAGTGGG 1159
DB 1145 GGAGACTGATGCTCCGATGATATCTCTTCAATGTTGAAGACCGCTGCTGATGAG 1204
QY 1160 CTATGATCTCTGCGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTTTCAGGGA 1219
DB 1205 CCTCAGAGACACGGGACGCTTGGCTCCGAGAGTTGGAAATTCATCTTCCATGCGGTTT 1264
QY 1220 TAAAGAGATCTTCTTGGGACACATCTCTGTTGTGTTT---AAACATGAGTTTCTTC 1275
DB 1265 TGGTCGAGATCTTCAAGCTGTGTAAATTCAGATCCTGTGGCAAGGATGGCAAAACACCC 1324
QY 1276 ATGATTGGGACCTATTTCAAGCTATGAGGAATAGGATCTGGCGGGGTTTGGTCCAGTTT 1335
DB 1325 ATGATTTCCAGGCTTTCGGATCTGAGGTTGGGATACGGCCGAGTTTCTGCTC-TATTAC 1383
QY 1336 GAAAGCGGGTTTATAGATCTCTCCGCTTGCATCAACGGAATAGAAAGAAATCAGCGG 1395
DB 1384 AACGTGTGTTTCAACGATCTGAGCTGATATCAATGGCTACGAGAGAGACAGCAT 1443
QY 1396 ATGTGCCCTGAAGAAATCTCAGAACTTCCACGTCGATGCCCATCTGAAAGTTAAAGGT 1455
DB 1444 CTTTCTATTTGTGGGGTTCACTTTTGCAGGCTCTGATGGCAATGAAATATTCAGAAA 1503
QY 1456 GTGTCTGTGAGCAGCAGCATATGCCATGTTCAAGTCAGGCGGATTCAGAAAGAAAGACA 1515
DB 1504 AGCCATGCGAAGACACGACTCTTTTGTGATCCCGTGTGTGGAATAGCAGAGAGGGGG 1563
QY 1516 AAAATAAAGTAAGCTTAAAGCGGGAATCTGAACCTTATGATAAAGTGTGTGTACA 1575
DB 1564 AGATTGAAGGTATGCTTGAACACGCGTCATCCGCTGTTTGAACAGGTCATCATATGGC 1623
QY 1576 TCTGAGCTGCCAAATATCCAACTCAGGCAATGGCTGATGACATGCAATATTTTTCAG 1635
DB 1624 GGCAGTGTGAGGCGGCTACAGTTGATTAACAGACTGGCCGGGATGAGACTTCTTCAGC 1683
QY 1636 GCACCAAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGATGCTCAAACTCTTCTG 1695
DB 1684 TACAAATATGAAACCCCGCTCGGAACTCGTCTGCGCTGTCAATTCAGACATCTTCATG 1743
QY 1696 ATGACTGAACGAAATTTGTGTTAGAACCATATCTCTCGCTTGTGTCTCATGACGGG 1755
DB 1744 GTCACGAAGCAAAAGTTTGGTTAACTCCGCACTCCAGCAGTATGATGACCGATGGG 1803
QY 1756 ATGCGAAAGCAATGATTTGCTGCTGACTATGATGCGAGATCCGAATGTTAAAGTCTA 1815
DB 1804 CTGTGCTGAGCTGTGTGATGATGATGAAATG-----CCAGCTGAGAGGGGCTT 1857
QY 1816 GTGCTCATCATGATATCATGAGAGAGCACTCCCAAGCTGTTGGCGGCTCCCGCAAA 1875
DB 1858 GTGCTTTTTCATCATGCTTTGATGATGATATCTATCCCGCGCATTCAGACATCATGACAG 1917
QY 1876 AAGAGCGATTAATGTGTGCTGCGGAGCAAAATTTGAGATCTTCCCGGCGTTGGCCAG 1935
DB 1918 AAGGAGCGGTGTGGAATTTGTCAAGGAGCTTGTGCTGCTTCTTGAATGCGCTTGT 1977
QY 1936 CACTATTTTCTGCTGCGCTGATTTAGACCAAAATGT 1973
DB 1978 CACTGTGCTCCAGTCAACGAAAGCACTAGCAATATGT 2015

```

RESULT 5
 US-08-458-831-8
 ; Sequence 8, Application US/08458831
 ; Patent No. 5824866

```

; GENERAL INFORMATION:
; APPLICANT: Slightom, Jerry L.
; APPLICANT: Tepfer, David A.
; TITLE OF INVENTION: R1 T-DNA Promoters
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GRAY, CARV, AMES & FRYE
; STREET: 401 B Street, Suite 1700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-4297
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,831
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,216
; FILING DATE: 25-JAN-1993
; APPLICATION NUMBER: US 06/725,368
; FILING DATE: 22-APR-1985
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhorst, Marile W.
; REGISTRATION NUMBER: 36,740
; REFERENCE/DOCKET NUMBER: P1020US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 699-2700
; TELEFAX: (619) 236-1048
; INFORMATION FOR SEQ. ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORGANISM: Agrobacterium rhizogenes
; STRAIN: Strain A4
; IMMEDIATE SOURCE:
; LIBRARY: Convolutus arvensis plant cells
; CLONE: Clone 7
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..2279
; OTHER INFORMATION: /label=ORF8
; OTHER INFORMATION: /note="Sequence ORF8 corresponds to bases 6609
; Patent No. 5824866
; OTHER INFORMATION: through 8888 of Seq. ID No. 5824866 19. "
; US-08-458-831-8

Query Match 6.0%; Score 136.4; DB 1; Length 2279;
Best Local Similarity 45.9%; Pred. No. 8.7e-37;
Matches 623; Conservative 0; Mismatches 721; Indels 14; Gaps 4;

QY 623 TCGACTTGCTCTACGACTGACAGCCGTTTTTTTGACCAATGTTCCATAGTGAAGCATG 682
DB 665 TTGACCTGCGGATACGACTATCAGTTTCTCCAGAGTGCATGCCATGATGTGTGG 724
QY 683 GCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGCGGCTCAATGGCGCTGGCATTTCCG 742
DB 725 GGCATTAATCCGAAGGCGACACATACCTAATGTTGCAATCGTAGAGCGGGGCTGTG 784
QY 743 GACTCGTGTGCAACGAACTGCTTCATGCTGGGTAGACGATGTTACAAATATATGAAG 802
DB 785 GCCTTGTGCTGCCACAGAACTACTTGGCGCTGGGTGCAAGAAATCACTCTTTTGATA 844
QY 803 CAAGTGATGTGTGGAGCA--GCTTGTCAATCTTTCAGGAGCGCTCTAGTG 859

```

Db	845	CCGTTGATGAGATCCGTAAGTTTGGGGGATCGCGATGCAAAAGCGACGTCACAG	904
Qy	860	TCTGTCGCGAAATGGGGCGATGCGATTTCTCTGTCGATCTGCTTTTCTTTC	919
Db	905	CCTTGACGTCGTTGGTGTCAATGCTTCTTCGCGAACCAACTTGTGCTGATCTATC	964
Qy	920	TGAGCGGTACGGCGCTGTCTTCGATGAGCGGTTCCAAATCCGGGCAAGTCGACTT	979
Db	965	TGATTAAGTTTAAATTCGCTTCAGCCTTCGTTTCTTGTGCGGCAACGACACAG	1024
Qy	980	ACTTGCTACCAAGGCGTCCAAATCATGTGAAAAGCCGGGACGTGCGACGAACTGT	1039
Db	1025	CACATATATTTCCGCGAAGAAAGCTATGCGATGCGACGCGGGCGAACGCTCCGCGGATAT	1088
Qy	1040	TCCATCCGCGTTTACACGCGTTGGCGCTGCGTTCTTGAAAGACGTTTTCATGACGATA	1099
Db	1085	TTTCAGCGGGTTCATGTCCGATGGAAGACATACTACTCAAGAGGTTGTAAACGAAATGACA	1144
Qy	1100	TTTGTTGGCTTCGCTGTGCTATTAATTCAGGCGTTGAAATCAGACACATTAAGTGGG	1155
Db	1145	GGAGACTGATGCTCCGATGATATCTTTTCAATGTAAAGAGCGTGTGCTGATGAAG	1204
Qy	1160	CTCATGACTCTGCGAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTTTCAGGGA	1219
Db	1205	CTTCAGAGACACGGGACGTTGGCTCGAGAGTTGGAATAATCTCTTTCATGCGCTTT	1264
Qy	1220	TAGAGAGATCTTCTGTGGCAACATCTCTGTGTG---AAACATGAGTTTCTCTC	1275
Db	1265	TGTCGAGATCTTCAGCTGTGTATTTGAGTCCGTGGTGGCAAGGCATGGCAACACCC	1324
Qy	1276	ATGATTTGGACCTTATCAAGCTAATGGAAATAGGATCTGGCGGGTTTGTTCAGTTT	1335
Db	1325	ATGATTTTCAGGCTTTCGGGATACGAGTTGGGATACGGCGAGTTTCGTC--TATTAC	1383
Qy	1336	GAAAGCGGGTTTATGAGATCTCCGCTGGTCATCAACGGATTAAGAAATACGCG	1395
Db	1384	AACGTGTTGTTTTCAGCATCTTGACTGATTAATCAATGGCTACGAGAGAGACGACAT	1443
Qy	1396	ATGTGCCCTGAAGAAATCTCAGAACTTCCACGTGCGATCGCATGGAAGTGAATACGCT	1455
Db	1444	CTTTCATTTGGTGGGTTCAACTTTTTCAGGCTGTGATGGCCATTGAATATTCAGAAA	1503
Qy	1456	GTGTCGTGAGCCAGCGCATATGCCATTTCAAGTCAGGCGGATTCAGAGGAAAAGACA	1515
Db	1504	AGCCATCGGAAAGACGACTCTGTTTGTATCCGTCGTCGAAATAGCCAAAGAGGCGGG	1563
Qy	1516	AAATTAAGATTAAGGCTTAAGAGCGGGAATATCTGAATTTATGATTAAGTGTGTCA	1575
Db	1564	AGATTTGAAGTATCTTTGAACACGCTCATTTGCGGTGTTTTTGACCAAGTATCATTTGGC	1623
Qy	1576	TCTGACTCGCAAAATATCAACTCAGGATTTGCTGACATCGATACCAATATTTTCAG	1638
Db	1624	GGCAGTGTGAGGCGCTACAGCTGATTAACAGACTGGCGGGGATGAGACTTCTTCAGC	1683
Qy	1636	GCACCACTGAACCAAGCGGTTGATTAACAGGCATATGACAGGATGTCAAAATCTTCTG	1695
Db	1684	TACAAATATCGAACCGCGCTGGAAAATCGTGTGCGCTGTCAATTCAGCACTCTTCATG	1743
Qy	1696	ATGACTGAACGAATTTCTGTTAGACCATATCTCTCCGTTTGTGTCTCATGACGGG	1755
Db	1744	GTCACGAAGCAAAAGTTTTGGGTTAATCTCGGCAATCCAGAGATGATATGACCGATGGG	1803
Qy	1756	ATCCCAAAAGCAGTATTTGCTGCGACTATAGTTCGACGATTCGAAATGTTAAAGTCTA	1815
Db	1804	CTTGTCCGTAGCTGTGTGTGATTTGACATCGAATG-----CCAGCTGAGAGAGGCTT	1857
Qy	1816	GTGCTCATAGTTATACATGGGAGACATCCACAAAGCTGTGGCGGTCGCCGACAA	1875
Db	1858	GTGCTTTTTCATATGCTTTTGGATGACATATATCCGCGGATCGACATCATGACAAAG	1917
Qy	1876	AAACAGGATTAATGTCTGCTGCGGGAGCGCAATTTTCGAGATCTTTTCCGGGCTTTGCCAG	1935
Db	1918	AAAGGACGGTCTTGAATTTGTGACGAGAGCTGTGCTGCTTTCCTGAACTGGCTTGT	1977

```

1936 CACCTATTTCCTCCGCGCTGATTTACGACCAAAATGT 1973
1937 ||||| ||||| ||||| ||||| ||||| |||||
1938 ||||| ||||| ||||| ||||| ||||| |||||
1939 ||||| ||||| ||||| ||||| ||||| |||||
1940 ||||| ||||| ||||| ||||| ||||| |||||
1941 ||||| ||||| ||||| ||||| ||||| |||||
1942 ||||| ||||| ||||| ||||| ||||| |||||
1943 ||||| ||||| ||||| ||||| ||||| |||||
1944 ||||| ||||| ||||| ||||| ||||| |||||
1945 ||||| ||||| ||||| ||||| ||||| |||||
1946 ||||| ||||| ||||| ||||| ||||| |||||
1947 ||||| ||||| ||||| ||||| ||||| |||||
1948 ||||| ||||| ||||| ||||| ||||| |||||
1949 ||||| ||||| ||||| ||||| ||||| |||||
1950 ||||| ||||| ||||| ||||| ||||| |||||
1951 ||||| ||||| ||||| ||||| ||||| |||||
1952 ||||| ||||| ||||| ||||| ||||| |||||
1953 ||||| ||||| ||||| ||||| ||||| |||||
1954 ||||| ||||| ||||| ||||| ||||| |||||
1955 ||||| ||||| ||||| ||||| ||||| |||||
1956 ||||| ||||| ||||| ||||| ||||| |||||
1957 ||||| ||||| ||||| ||||| ||||| |||||
1958 ||||| ||||| ||||| ||||| ||||| |||||
1959 ||||| ||||| ||||| ||||| ||||| |||||
1960 ||||| ||||| ||||| ||||| ||||| |||||
1961 ||||| ||||| ||||| ||||| ||||| |||||
1962 ||||| ||||| ||||| ||||| ||||| |||||
1963 ||||| ||||| ||||| ||||| ||||| |||||
1964 ||||| ||||| ||||| ||||| ||||| |||||
1965 ||||| ||||| ||||| ||||| ||||| |||||
1966 ||||| ||||| ||||| ||||| ||||| |||||
1967 ||||| ||||| ||||| ||||| ||||| |||||
1968 ||||| ||||| ||||| ||||| ||||| |||||
1969 ||||| ||||| ||||| ||||| ||||| |||||
1970 ||||| ||||| ||||| ||||| ||||| |||||
1971 ||||| ||||| ||||| ||||| ||||| |||||
1972 ||||| ||||| ||||| ||||| ||||| |||||
1973 ||||| ||||| ||||| ||||| ||||| |||||
1974 ||||| ||||| ||||| ||||| ||||| |||||
1975 ||||| ||||| ||||| ||||| ||||| |||||
1976 ||||| ||||| ||||| ||||| ||||| |||||
1977 ||||| ||||| ||||| ||||| ||||| |||||
1978 CACTGTGCTCCAGTCAACGAAACTACGAACATATGT 2015

```

```
LOCATION: complement (4607..4918)
OTHER INFORMATION: /label= ORF5SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5143..6216
OTHER INFORMATION: /label= ORF6SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (5071..5643)
OTHER INFORMATION: /label= ORF7SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6609..8888
OTHER INFORMATION: /label= ORF8SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (6576..6830)
OTHER INFORMATION: /label= ORF9SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9748..10044
OTHER INFORMATION: /label= ORF10SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (10509..11282)
OTHER INFORMATION: /label= ORF11SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12465..13002
OTHER INFORMATION: /label= ORF12SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 13723..14319
OTHER INFORMATION: /label= ORF13SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15659..16210
OTHER INFORMATION: /label= ORF14SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (16517..17545)
OTHER INFORMATION: /label= ORF15SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (17737..18189)
OTHER INFORMATION: /label= ORF16SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (18177..18743)
OTHER INFORMATION: /label= ORF17SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (19031..19390)
OTHER INFORMATION: /label= ORF18SUBSEQUENCE
US-08-008-216-19

Query Match 6.0%; Score 136.4; DB 1; Length 21126;
Best Local Similarity 45.9%; Pred. No. 4.8e-36;
Matches 623; Conservative 0; Mismatches 721; Indels 14; Gaps 4;

QY 623 TCAGACTGCTCTACGACTGCGAGCCGTTTTCACATGTTCCGATAGTGAAGCGATCG 682
DB 7273 TTGAACCTGCGTAGACTATACAGTTTCTCCAGAGTGCAGTCCCATGATGCTGG 7332
QY 683 GCTTCTTTCGAGAGATGTTCTTAAGCCGAAAGTGGCGTCAATTGGCGCTGGCATTTCCG 742
DB 7333 GGCAATTATCCGAAAGGCGACCACTAGTGTTCATCGATGATGAGGCGGCTGTCTG 7392
QY 743 GACTCGGTGGGCAAGAACTGTTTATGCTGGGGTGAAGAGATGTTACATATATAG 802
DB 7393 GCCTTGTGTGCGCAAGAACTACTTGGCGCTGGCGTCAAGAAATACCTTTTCGATA 7452
QY 803 CAAGTATCGTGTGAGAGCAA---GCTTGTGCATATGCTTTACGGAGCGCTCTAGTG 859
```

```
DB 7453 CCGTATAGATCCGTAATTTTGGGGCATCCGCCATGCCAAACGGCGACGCTCACAGG 7512
QY 860 TCGTGGCCGAATATGGGGCGATGCGATTTCTCTCTGCTGCATTTGCTTTCTTCC 919
DB 7513 CTTGACGTGCTTGGGTGTCATGCTTTCTCGCCCAACCACTTGTGCTATCTATC 7572
QY 920 TCAGAGCTTACGAGCCCTGCTTCGATGAGGCGCTTCCCAATCCGGCAGAGTGCACATT 979
DB 7573 TGGATTAAGTTTAAATTCGCTCCAGCTTCGTTTCTGTGCGGCAACGACACAG 7632
QY 980 ACTTGTCTACCAAGGCGTCAATACATGTGAAACCGGGCAGCTGCCACCGAAGCTGT 1039
DB 7633 CACTATATTTCCGCCAGAAACCTACGATGACACGCGGGCAAGCTCCGCCGGGATAT 7692
QY 1040 TCCATGCGCTTTAAACAAGGTTGCGCGTCTTTGAAGAACGTTTTCATGAGCGAGATA 1099
DB 7693 TTCAGCGGTACATGTGGATGGAACACTCTTACCAAGGGTGTGAACGAATGGCA 7752
QY 1100 TTGTGTGGCTTCGCTGCTGCTTACTCAGGCTTGAATCAGACACATTAGGTGG 1159
DB 7753 GGAAGACTGATGGCTCCGATGATATCTTTGATGTTAAAGCGCTCGTGTGATGAAG 7812
QY 1160 CTCATGACTCTTGGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTTCAGGA 1219
DB 7813 CTCAGAACACAGCGCAGCTTGGCTCCGAGAGTTCGAAATTCATTCCATCCCGTTT 7872
QY 1220 TAGAGAGATCTTTCTGGGCAACATCTCTGCTGTGCT---AAATGAGTTCCTC 1275
DB 7873 TGGTCGAGATCTTACGCTGTGTAATTCAGTCTGTGTGGCAAGGATGCAAAACCCC 7932
QY 1276 ATGATTGGACCTTATTCAGCTATATGGAATAGATCTGCGGGGTTTGGTCCAGTTT 1335
DB 7933 ATGATTTTCAGAGGTTTTCGGGATATCTAGTTGGGATACGGCCAGTTTGTCTC-TATTC 7991
QY 1336 GAAAGGGGTTTATAGATCTCTCGCTTGTTCATCAACGATATGAAGAAATCACGGG 1395
DB 7992 AAGGTGTGTTTCAACGATCTGGAATGATGATTAATGATGATGACAGAGGACAGCAT 8051
QY 1396 ATGTGCCCTGAAGGAATCTCAGAATCTTCCAGCTGGATGCGCATCTGATGCTTAAGGT 1455
DB 8052 CTTTCTATTGTGGGGTCAACTTTTTCAGGCTGTGATGCGCATTAATTAATTCAGAAA 8111
QY 1456 GTGTCTGTAGGCGGCAATATGCAATGTTCAAGTCAGGCGCATTCAGAAAGAAAGACA 1515
DB 8112 AGCATGCGAAAGCAAGACTCTTTTGTATCCGCTCGTGGAATTAACCAAGAGGGCGGG 8171
QY 1516 AAAATTAAGATTAAGGCTTAAGACGGGATATCTGAACTTTATGATTAAGTGTGTACA 1575
DB 8172 AGATTGAAGGTATGTTGAAACACGCTCATTCGCTGTTTGAACAGGTCATCATTTGCG 8231
QY 1576 TCTGACTGCGAAATATCCAACTCAGGCAATGCTGTGATGCAATGCAATATTTTCAG 1635
DB 8232 GGCAGTGTGAGGCGCTACAGTTGATTAACAGCTGCGGGATAGATGATTCCTTCAGC 8291
QY 1636 GCACAGTGAACCAAGCGGTTGATTAACAGCATATGACAGATGTCGAAACTCTTCTG 1695
DB 8292 TACAAATATGAACCCCGCTCGGAAACTGCTGCTCGCTGTCAATTCAGACCTTTCATG 8351
QY 1696 ATGACTGAACGAAATATCTGTTTGAACCAATATCTCCGCTTTGTGCTCATGACGGG 1755
DB 8352 GTCACGAGCAAAAGTTTGGTTAACTCCGGATCCACAGCATGATGATGACGATGGG 8411
QY 1756 ATGCAAAAGCAATGATTTGCTGAGCTATGATGCGCAGATTCGCAATGTTAAAGCTTA 1815
DB 8412 CTGTGCTGAGTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8465
QY 1816 GTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1875
DB 8466 GTGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8525
QY 1876 AAAGACGATTAATGCTGTGCGGGAACGAATTTGAGATCTTTCGCGGGTTCGCGAG 1935
```


Db	7293	GCCTGTGTCGCCACAGAACTACTTGGCGCTGGGGTCCAAAGAAATACCTCTTTTCGATA	7452
Oy	803	CAAGTATCGTGTGGAAGCAA---GCTTTGTCATACGCTTTAAGGAGCCTCTTACTG	859
Db	7453	CCGTTGATGAAATCCGTAGTTTGGGGGATGCCGATGCCAAAGCGGACCTCACAAG	7512
Oy	860	TCGTGGCGAAATGGGGGCGATGGGATTTCTCCGCTGCATTCGCTGTGTTTTCTTCC	919
Db	7513	CTTGAGCTGCTGGTGTCAATGCTTTCTCCGCAACCACTTGGCTGTCAATATC	7572
Oy	920	TCGAGCGTTAAGGCGCTGTCTTCGATGAGGCCGTTCCAAATCCGCGCACAGTCGACTT	979
Db	7573	TGGATAGATTAGAAATCCGTCACGCTTGCTTTTCTTGTCGGGACGACACACAG	7632
Oy	980	ACTTGCTACCAAGCGCTCCAATACATGTGAAAGCCGGCAGCTGCACCGAACTGT	1039
Db	7633	CACATATATTTCCGCGAGAAACGCTACGCATGGCAGCGCGGCGAAGCTCCGCGGGATAT	7692
Oy	1040	TCCATTCGGGTTTACAAAGGTTGGGGTGCTCTTGAAGAGAGGTTTCAATGAGGAGATA	1099
Db	7693	TTCAAGCGGGTACATGTCCGATGAAAGACATACTTACAAAGGGTGTGACGAAATGGCA	7752
Oy	1100	TTGTGTTGGCTTCGCTCTGCTGCTATTTACTACAGCCTTGAATAACAGACATTAAGTGGG	1159
Db	7753	GGAAACGATAGGCTCCGATGGAATATCTTTCAATGTTGAAAGAGCGTGTGCTATGAAG	7812
Oy	1160	CTCATGACTCTGCGCAATTTGGCTGAACCGTTTCGGAGGGAAGTCTTCTTCAAGGA	1219
Db	7813	CCTCAGAAAGCAGGAGCTTTGGCTCCGAGAGTTCCGAAATTCACATTTTCATGCCGTTT	7872
Oy	1220	TAGAGAGATCTTTCTGGGCAACATCTCTGCTGGT---AAACATGAGATTTTCTC	1275
Db	7873	TGCTCGAATCTTACGCTGTGATATTCGATCTCCGTGGGCAAGGCAATGGCAACCCC	7932
Oy	1276	ATGATTTGGGACCTATTCAAAGCTAATGGGAATAGATCTGGCGGGGTTTGTCAGATTTT	1335
Db	7933	ATGATTTTGAAGCTTTGGGGAATCTGAGAGTTGGGATACGCGAGTTTCGTC--TATTAC	7991
Oy	1336	GAAAGCGGGTTTATTGAGATCTCTCGCTTGTGTCATCAACGGATATGAAGAAATCAAGCG	1395
Db	7992	AACGTGTTGTTTTCACAGATCTTGGACTGATTTATCAATGGCTACGAGAGGACAGCAT	8051
Oy	1396	ATGTGCTTGAAGGAATTCAGAACTTCCACAGTCGGAATCGCATCTGAAGTGGTTAACGT	1455
Db	8052	CTTTCTATTTGGTGGGTTCAACTTTTGGAGCTCGATGCGCATTGAATTTTCCAGAA	8111
Oy	1456	GTCGTCTGAGCCAGCGCATATGCCATGTTCAAGTCAGGGGCAATTCAGAGAAAGAACCA	1515
Db	8112	AGCCATGCGAAAGCACACACTCTGTTTGTATCCCTGCGTGCATATGCCAAGAGAGGCGGG	8171
Oy	1516	AAAAATAAAGATAGGCTTTAAGAGCGGGAATCTGAACCTTATATGATAGGTGGTGTACA	1575
Db	8172	AGATTGAAGGATGTGTGAACACGGGTCAATTCGGGTGTTTTTGAACAAGTCATCATTTGGC	8231
Oy	1576	TCTGGACTTCGCAATATATCAACTCAGGCAATGCTGTGACATGCGATACCAATATTTTACG	1635
Db	8232	GCGAGTGTGAGCGCGCTACAGTTGATTAACAGACTGCGCGGGGATGAGACTTCTTTCAGC	8291
Oy	1636	GCACAGTGAACAAAGCGGTGATTAAGACCAATATGACAGATGCTCAAACTCTTCTG	1695
Db	8292	TACAATATCGAACCCGCGTCGAAACCTGCTGCGCTGTCAATTCACACTCTTCAATG	8351
Oy	1696	ATGACTGAACAAAATTCCTGTTAGACCATATCTCCGCTTGTGTCCTCATGAGCGGG	1755
Db	8352	GTCACGAAGCAAAAGTTTGGTTTAATCTCCGCAATCCAGCAGATATATGAGACGATGGG	8411
Oy	1756	ATCGCAAAAGCAGTGTATTCCTGGACTATAGTCGACAGATCCGAATGTAAAGTCTTA	1815
Db	8412	CTTGTCGCTGAGCTGTGTGATTCATGACATCGAATG-----CCAGCTGAGAGAGGCGCTT	8465
Oy	1816	GTCGCTATCAGTTATATCAATGGGAGACGATCCCAAGCTGTGGCGGCTCCCGACAAA	1875
Db	8466	GTCGTTTTTCACTAATGCTTTGGATATACATCTATCCGCGCGATGAGCATCATATGACAG	8525

[illegible]

```

1      RESULT 8
2      OS-458-831-19
3      Sequence 19, Application US/08458831
4      Patent No. 582486
5      GENERAL INFORMATION:
6      APPLICANT: slighthom, Jerry L.
7      APPLICANT: Tepfer, David A.
8      TITLE OF INVENTION: R1 T-DNA Promoters
9      NUMBER OF SEQUENCES: 19
10     CORRESPONDENCE ADDRESSES:
11     ADDRESSSEE: GRAY, CARY, AMES & FRYE
12     STREET: 401 B Street, Suite 1700
13     CITY: San Diego
14     STATE: California
15     COUNTRY: USA
16     ZIP: 92101-4297
17     COMPUTER READABLE FORM:
18     MEDIUM TYPE: Floppy disk
19     COMPUTER: IBM PC compatible
20     OPERATING SYSTEM: PC-DOS/MS-DOS
21     SOFTWARE: PatentIn Release #1.0, Version #
22     CURRENT APPLICATION DATA:
23     APPLICATION NUMBER: US/08/458,831
24     FILING DATE: 02-JUN-1995
25     CLASSIFICATION: 800
26     PRIOR APPLICATION DATA:
27     APPLICATION NUMBER: US 08/008,216
28     FILING DATE: 25-JAN-1993
29     APPLICATION NUMBER: US 06/725,368
30     FILING DATE: 22-APR-1985
31     ATTORNEY/AGENT INFORMATION:
32     NAME: Barnhorst, Marnie W.
33     REGISTRATION NUMBER: 36,740
34     REFERENCE/DOCKET NUMBER: P10200U1
35     TELECOMMUNICATION INFORMATION:
36     TELEPHONE: (619) 236-1048
37     TELEFAX: (619) 236-1048
38     INFORMATION FOR SEQ ID NO: 19:
39     SEQUENCE CHARACTERISTICS:
40     LENGTH: 21126 base pairs
41     TYPE: nucleic acid
42     STRANDEDNESS: single
43     TOPOLOGY: linear
44     MOLECULE TYPE: DNA (genomic)
45     HYPOTHEetical: NO
46     ORIGINAL SOURCE:
47     ORGANISM: Agrobacterium rhizogenes
48     STRAIN: STRAIN A4
49     IMMEDIATE SOURCE:
50     LIBRARY: CONVOLVULUS ARVENSIS PLANT CELLS
51     CLONE: CLONE 7
52     FEATURE:
53     NAME/KEY: misc feature
54     LOCATION: complement (937..2262)
55     OTHER INFORMATION: //label= ORF1SUBSEQUENCE
56     FEATURE:
57     NAME/KEY: misc feature
58     LOCATION: complement (2649..3458)
59     OTHER INFORMATION: //label= ORF2SUBSEQUENCE
60     FEATURE:
61     NAME/KEY: misc feature
62     LOCATION: 3726..4799
63     OTHER INFORMATION: //label= ORF3SUBSEQUENCE

```



```

FEATURE:
NAME/KEY: misc feature
LOCATION: complement (4041..4400)
OTHER INFORMATION: /label= ORF4SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (4607..4918)
OTHER INFORMATION: /label= ORF5SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 5143..6216
OTHER INFORMATION: /label= ORF6SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (5071..5643)
OTHER INFORMATION: /label= ORF7SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 6609_8888
OTHER INFORMATION: /label= ORF8SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (6576..6830)
OTHER INFORMATION: /label= ORF9SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 9748..10044
OTHER INFORMATION: /label= ORF10SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (10509..11282)
OTHER INFORMATION: /label= ORF11SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 12466..13002
OTHER INFORMATION: /label= ORF12SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 13723..14319
OTHER INFORMATION: /label= ORF13SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: 15659..16210
OTHER INFORMATION: /label= ORF14SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (16517..17545)
OTHER INFORMATION: /label= ORF15SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (17737..18189)
OTHER INFORMATION: /label= ORF16SUBSEQUENCE
FEATURE:
NAME/KEY: misc feature
LOCATION: complement (19031..19399)
OTHER INFORMATION: /label= ORF18SUBSEQUENCE
;S-08-458-831..19

```

Query March	6.0%	Score 136.4	DB 1	Length 2126
Best Local Similarity	45.9%	Pred. 0.48e-36		
Matches 623	Conservative	0	Mismatches 721	Indels 14
				Gaps 4

QY	623	TCGACTTGCTTACGACTGCAGACCGTTTTTTTGACCAATGTTCCGATAGTGACGAGATCG	682
Db	7273	TTGACCTCGGGTAGGACTATCAGGTTTTTCTCCAGAGATGGGATGCCATGATGGTGTGG	7333
QY	683	GCTTCTTTCCGAGAGATGTTCTTAAGCCGAAAGTGCGCGTCATTTGCGCCTGGCATTTTCG	742
Db	7333	GGCATTTATCCGAAAGCGCACCTACTTCACTTAGTGTTCATGATAGAGACGGCGGTCTTG	7392

OY	743	GACTCGTGATGGCAAAACGAACTGCTTCAATGCTGGGGTATGAGATGTTCAATATATGAA	802
Db	7393	GCCTTGTTGCTGCCAAGAACTACTTGGCCGCTGGCTCAAGAAATCACTCTTTTCGATA	7455
OY	803	CAAGTGAATCGTGTGGAGGCA---GCTTGTGTCACATGCTTTAGAGGACGCTCTAAGT	859
Db	7453	CCGTTGATGAGATCCGATGTTTGGGGGCATCGCGATGCGAAAGGAGCACCTCACAG	7512
OY	860	TGCGTGGCGAAATGGGGGCGATGGGATTTCTCTCGTGCAATCTGCGTGTCTTTCTTCC	919
Db	7513	CTTTGACGTGCTTGGGTGTCATGCTTCTTCCGCAACAACTTGGCTGTGCAATGATC	7572
OY	920	TCGAGCGTTACGGCGCTGTCTTCGATGAGGCGGTTCCCAATCCCGGCAAGTGCACACTT	979
Db	7573	TGGAATAGATTAGAAATTCGTCGACGCTTGTTTCTTGTCGGGCAACGACACACAG	7633
OY	980	ACTTGCTCTACCAAGGCGTCCAAATCATGTGAAAGCCGGGCACTGCCACCGAACTGT	1033
Db	7633	CACATATATTTCCGCGCAAGAAAGCTACGATGCGACGGGGCAAGCTCCGCGGGGATAT	7692
OY	1040	TCCATCGCGTTTACACGCGTGGCGTGCGGCTTCTGAAGACGATTTTCATAGCGAGATA	1099
Db	7693	TTTACCGGGTACATGTGGGATGGAAGACATCTCTACCAAGGGTGTCAACGGAATGCC	7755
OY	1100	TTGTGTGGCTTGCCCTGTGCTATTACTCAGGCGTTGAAATCAGACACATTAAGTGGG	1155
Db	7753	GGAGACTGATGGCTCCGATGGATATCTTTCAATGTGAAGACGTCGTGATGTAAG	7812
OY	1160	CTCATGATCTCGTGCCAATTTGGCTGAACGTTTCCGGAGGGAGTCTTCTCTCAGGGA	1219
Db	7813	CCTCAGAAAGCAGGCACTGTGGCTCCGAGAGTTGGAATAATCATCTTCCATGCGCTTT	7872
OY	1220	TAGAGAGATCTTCTCTGGGCAACATCTCTCTGTGGT---AAACATGAGTCTTCTCTC	1275
Db	7873	TGCTGAGATCTTCAAGCTGTGGTAAATTCAGATCCTGGTGGCAAGCATGGCAACCC	7932
OY	1276	ATGATTTGGGACTTATTCAGCTAATGGGAATAGATCTGGCGGGGTTTGGTCCAGTTT	1335
Db	7933	ATGATTTGAGGCTTTGGGATACGATGGTGGATACGGCCGAGTTTCGTCC-TATTAC	7992
OY	1336	GAAAGCGGGTTTATTAGATCTCCGCTGGTCACTCAACGGAATGAAAGAAATCAGCGG	1399
Db	7992	AACGTGTTGTTTCAAGATCTTGGACTGATTTCAATGGCTACAGAGAGACACAGAT	8053
OY	1396	ATGTGCCCTGAAGGAATCTCAGAACTTTCAGCTGGATCCGATCTGAAGTGTAAACGT	1455
Db	8052	CTTCTCTATGTTGGGGTTCAACTTTTCAGGCTCTGATGGCATTTGAATAATTTCCAGAA	8111
OY	1456	GTCGCTGTGAGCCGACGCAATGTCAGTTCAGTCAACGCGGATTCAGAAAGAAAGCA	1515
Db	8112	AGCCATCGAAAGCAGCACTCTGTTTGAATCCGTCGTGAATAAGCCAAAGAGAGCGGG	8171
OY	1516	AAATATAAGATTAAGGCTTAAAGCGGGATATCTGMACTTTATGATTAAGTGTGTGACA	1575
Db	8172	AGATTTGAAGGATATGCTTGAAGAACGTCATTTCCGGTGTTTTGAACAGGTATCATTTGC	8231
OY	1576	TTGGAAGCTCGCAATATCAACACTCAGCAATTGGCTTGCATATGCGATACCAATATTTTTCAG	1635
Db	8232	GGCAGTCTAGAGCGCGCTACAGTTGATTAACAGATGCGCCGGGGATGAGACTTCTTCAGC	8292
OY	1636	GCACCAAGTGAACCAAGCGGTGATTAACAGCAATATGACAGATGCTCAAAATCTTCTGTG	1695
Db	8292	TACAAATATCGAACCCGCGTCGAAACTCGTCTGCGCTGTCAATTAACGACTCTTCAATG	8351
OY	1696	ATGACTGAACGAAATTTCTGGTTAGACATATCTCCGCTTGTGTCTTCATGAGACGGG	1755
Db	8352	GTCACGAAGCAAAAGTTTGGGTTAACTCCGGCATCCACAGACGATATGACCGATGGG	8411
OY	1756	ATGCGAAAGCAGTATGCTCGGACTATAGAGTGGCAGAGATCCGAATGTGAAGGTCTA	1815
Db	8412	CTTGTCCGTAGCGTGTGTGCATTTGACATCGAATCG-----CCAGCTGAGAGAGGGCTTT	8465

SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,037
FILING DATE: June 13, 1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/005,167
FILING DATE: October 13, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 220/156
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1642 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 97...1446
US-08-665-037-1

Query Match 1.6%; Score 36; DB 2; Length 1642;
Best Local Similarity 62.0%; Freq. No. 0.15;
Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1814 TAGTCTCATCATGATTATACATGGAGAGACACTCCCAAGCTGTGGCGGTCCCGACA 1873
DB 728 TGTGGCCAGCATCAAGCAGCAGGAGAGACGAGCCCGCTGCTGCTGTGACCCCGACA 787
QY 1874 AAAAAGCGGATTATGTCTGCTCGGAGACCA 1905
DB 788 CAGATGAACACTTCAAGCGGCTTCGGGTCA 819

RESULT 13
US-08-666-067-1
Sequence 1, Application US/08666067
Patent No. 5922842
GENERAL INFORMATION:
APPLICANT: Seedorf, Klaus
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT
TITLE OF INVENTION: OF TKA-1 RELATED
TITLE OF INVENTION: DISORDERS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSES: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/666,067
FILING DATE: June 13, 1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/005,421
FILING DATE: October 13, 1995
ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 220/157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1642 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 97...1446
US-08-666-067-1

Query Match 1.6%; Score 36; DB 2; Length 1642;
Best Local Similarity 62.0%; Freq. No. 0.15;
Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1814 TAGTCTCATCATGATTATACATGGAGAGACACTCCCAAGCTGTGGCGGTCCCGACA 1873
DB 728 TGTGGCCAGCATCAAGCAGCAGGAGAGACGAGCCCGCTGCTGCTGTGACCCCGACA 787
QY 1874 AAAAAGCGGATTATGTCTGCTCGGAGACCA 1905
DB 788 CAGATGAACACTTCAAGCGGCTTCGGGTCA 819

RESULT 14
US-08-732-870-1
Sequence 1, Application US/08732870
Patent No. 5945523
GENERAL INFORMATION:
APPLICANT: Seedorf, Axel
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT
TITLE OF INVENTION: OF TKA-1 RELATED
TITLE OF INVENTION: DISORDERS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSES: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/732,870
FILING DATE: October 15, 1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/666,037
FILING DATE: June 13, 1996
APPLICATION NUMBER: 08/666,067
FILING DATE: June 13, 1996
APPLICATION NUMBER: 60/005,167
FILING DATE: October 13, 1995
APPLICATION NUMBER: 60/005,421
FILING DATE: October 13, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 222/247

TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1642 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 97...1446
US-08-732-870-1

Query Match 1.6%; Score 36; DB 2; Length 1642;
Best Local Similarity 62.0%; Pred. No. 0.15;
Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1814 TAGTGCTCATGATTATATGAGAGAGACCTCCCAAGCTGTGGCGTCCCGACA 1873
DB 728 TGGTGGCCAGCATCAAGCAGCGAGAGACAGGCCCGCTGCTGTGACCCGACA 787

QY 1874 AAAAGAGCGATTATGTCTGCTCGGAGCGCA 1905
DB 788 CAGATGACACTTCAAGCGCTTCGGGTACA 819

RESULT 15
US-09-415-277C-4/C
Sequence 4, Application US/09415277C
Patent No. 6531308
GENERAL INFORMATION:
APPLICANT: Hershenberger, Charles
APPLICANT: Payson, Robert
TITLE OF INVENTION: Ketoreductase Gene and Protein from Yeast
FILE REFERENCE: X-11325A
CURRENT APPLICATION NUMBER: US/09/415,277C
CURRENT FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: US 09/182,985
PRIOR FILING DATE: 1998-10-30
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1032
TYPE: DNA
ORGANISM: S. cerevisiae
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1032)
OTHER INFORMATION:
US-09-415-277C-4

Query Match 1.6%; Score 35.2; DB 4; Length 1032;
Best Local Similarity 53.7%; Pred. No. 0.2;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1498 ATTCAGAGGAAAGACAAATAAAGATTAGGCTTAAGACGGGATATCTGAATTAT 1557
DB 624 ATCGAAAGTTGAGAGCAAAAGAGAGAGGCTGACGTGTGATTGATTGAT 565

QY 1558 GATTAAGTGTGTCATCATCTGAGCTGGCAATATCCAACTCAGAGCATTGCGTCAGATGC 1617
DB 564 GTGATCTTCATCTCTTTTGTGAAGTCCAGGAGCCCTTTTACCAAACTTCTTGATGC 505

QY 1618 GATACCAATATTTTC 1633
DB 504 AAGTAAGCAATTATC 489

Search completed: November 23, 2003, 21:26:57
Job time: 147.658 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 21:01:51 ; Search time 4831.41 Seconds
(without alignments)
11409.196 Million cell updates/sec

Title: US-09-434-837-10

Perfect score: 2268
Sequence: 1 actgcagctcaccctctcct.....actaccgcagtagaattag 2268

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST: *
1: em_estda: *
2: em_esthum: *
3: em_estin: *
4: em_estnu: *
5: em_estcov: *
6: em_estpl: *
7: em_estro: *
8: em_estc: *
9: gb_estc1: *
10: gb_estc2: *
11: gb_estc3: *
12: gb_estc4: *
13: gb_estc5: *
14: gb_estfun: *
15: em_estfun: *
16: em_estcom: *
17: em_gss_hum: *
18: em_gss_inv: *
19: em_gss_pln: *
20: em_gss_vrt: *
21: em_gss_fun: *
22: em_gss_man: *
23: em_gss_mus: *
24: em_gss_pro: *
25: em_gss_rtd: *
26: em_gss_phg: *
27: em_gss_vrl: *
28: gb_gss1: *
29: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51.6	2.3	1201	13	BX381961 BX381961
2	49.8	2.2	618	9	AU239500 AU239500
3	44.8	2.0	564	12	BJ123611 BJ123611
4	44.2	1.9	461	12	BJ101485 BJ101485

	5	44.2	1.9	486	12	BJ103567
	6	44.2	1.9	508	9	AU199996
	7	44.2	1.9	511	12	BJ108542
	8	44.2	1.9	513	9	AU203802
	9	44.2	1.9	529	12	BJ120847
	10	44.2	1.9	540	9	AU201678
	11	44.2	1.9	560	14	CB401025
	12	44.2	1.9	584	12	BJ121613
	13	44.2	1.9	586	12	BJ121635
	14	44.2	1.9	597	9	AU207635
	15	44.2	1.9	603	12	BJ101145
	16	44.2	1.9	608	9	AU204573
	17	44.2	1.9	619	12	BJ117235
	18	44.2	1.9	634	12	BJ124022
	19	44.2	1.9	655	12	BJ103250
	20	43.8	1.9	1201	13	BX381961
	21	43.4	1.9	482	12	BJ101110
	22	43	1.9	38	9	AU109666
	23	42.6	1.9	631	12	BJ117453
	24	42.4	1.9	850	13	BU348455
	25	42	1.9	362	14	CA926583
	26	42	1.9	1201	13	BX376097
	27	40.8	1.8	999	29	CNS0601H
	28	40.6	1.8	310	13	BU821300
	29	40.6	1.8	411	9	A1102326
	30	40.6	1.8	479	9	CA934750
	31	40.6	1.8	1376	29	AC376243
	32	39.4	1.7	284	9	A1101650
	33	39.4	1.7	456	9	AA943462
	34	39.4	1.7	462	12	BI300778
	35	39.4	1.7	511	10	BC378762
	36	39.4	1.7	560	12	BI296551
	37	39.2	1.7	404	12	BJ305833
	38	39.2	1.7	410	12	BI611075
	39	39.2	1.7	437	13	BY467572
	40	39.2	1.7	501	14	CA614930
	41	39.2	1.7	502	12	BI368223
	42	39.2	1.7	554	9	A1457056
	43	39.2	1.7	564	12	BI566474
	44	39.2	1.7	622	12	BI627861
	45	39.2	1.7	623	12	BI588865

ALIGNMENTS

RESULT 1
LOCUS BX381961 1201 bp mRNA linear EST 08-MAY-2003
DEFINITION BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Clone CS0D1072XF05 3-PRIME, mRNA sequence.

ACCESSION BX381961
VERSION BX381961.1 GI:30453007

KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifestech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope Sequence ID: CS0D1072C03NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"

FEATURES
source

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D10721F05"
/issue="PLACENTA COT 25-NORMALIZED"
/clone_1db="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-cloned (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      95 a      191 c      115 g      55 t      745 others
ORIGIN

```

```

Query Match      2.3%; Score 51.6; DB 13; Length 1201;
Best Local Similarity 5.9%; Pred. No. 0.0025;
Matches 49; Conservative 288; Mismatches 488; Indels 2; Gaps 1;

Qy 1217 GGATAGAGAGATCTTTCTGGGACACATCTCTCTGGTGTAAACATGAGATTTCTCA 1276
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 351 GGNANKKAAANNKKKACNNCCNMAKCCNNACNNKKKKKACNANNNNKKCMM 410

Qy 1277 TGATTGGGACCTATTCAAGCTAATGGGAATAGATCTGGGGGTTTGTCAGTTTG 1336
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 411 NNNKKKKKCNKKKKKKKKKKKKKKKKKKKKNNNNNNNNNNNNNNNNNNNTK 470

Qy 1337 AAAGCGGTTTATGATCTCTCGCTTGTCATCAAGATATGAG--AAATCAGC 1394
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 471 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 530

Qy 1395 GATGTGCGCTGAAGAATCTCAGAACTTCACGTCGATCGCATCTGAAGTGTAAAG 1454
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 531 CNMNMNMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 590

Qy 1455 TGTGTCTGTAGCGGCGCATATGCCATGTTCAAAGGCGCATTCAGAGAAAAGAC 1514
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 591 KNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 650

Qy 1515 AAAAATAAAGTAAAGCTTAAGCGGGATATCTGAATTAAGTATGATGATGATG 1574
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 651 MNKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 710

Qy 1575 ATCTGAGCTGCAAAATATCCAACTCAGCACTTGCCTGATGATGATATTTTCA 1634
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 711 KTKKKKKKNNKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 770

Qy 1635 GGCACCACTGAACCAAGCGGTGATTAAGCCATATGACAGATCGTCAAACTCTTCT 1694
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 771 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 830

Qy 1695 GATGACCTGAAGAAATTTGTGTTAGACATATCTCCGCTGTGTCTCATGAGCG 1754
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 831 MKBMMKCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCK 890

Qy 1755 GATGCAAAAGCAAGTATTTGCTGAGCTATGATGCTCAGAGATCGAATGGTAAAGCT 1814
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 891 MMBBKKKMMKKKMMKKKMMKKKMMKKKMMKKKMMKKKMMKKKMMKKKMMKK 950

Qy 1815 AGTGTCTCATGATTATCATGAGAGAGCACTCCCAAGCTGTGGCGGTCCCCGCAA 1874
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 951 CVMKMKMKMCKCKVMKMKMKCAKKKKMKCMVKKMCKMCKMCKMCKMCKMCKM 1010

Qy 1875 AAAAGAGCGATTATGTCTGCTGGCGGAGCGAATTCAGATCTTCCCGCGTTTGCCA 1934
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1011 MNNNNCAKMMMBMMMBMMMBMMMBMMMBMMMBMMMBMMMBMMMBMMMBMM 1070

Qy 1935 GCACTATTTCCTGCTGCGGTGATTAAGACCAAAATGTTATCAATGATTTGGCTTAC 1994
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1071 VCKMBBCKMCKGCMCKGCMCKGCMCKGCMCKGCMCKGCMCKGCMCKGCMCK 1130

Qy 1995 AGACGAAATGCGCGGAGGCTTTCAAATCAACCGCGGTGTGAGG 2041
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1131 BKVCVKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 1177

```

```

RESULT 2
AU239500      618 bp      mRNA      linear      EST 01-APR-2002
LOCUS
DEFINITION
AU239500 RAFL19 Arabidopsis thaliana cDNA clone RAFL19-89-D06 5',
mRNA sequence.
ACCESSION
AU239500
VERSION
AU239500.1 GI:19878669
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 618)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shingawa,A., Muramatsu
,M., Hayashizaki,Y. and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA
Unpublished
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@rc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pluscript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/planc/index_e.html) for further
details.
FEATURES
    source
        1..618
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /db_xref="taxon:3702"
            /clone="RAFL19-89-D06"
            /issue_type="mixture of silique and flower"
            /lab_host="DH10B"
            /clone_1db="RAFL19"
            /note="Site 1: BamHI; Site 2: SalI; Subtraction Library"
BASE COUNT      158 a      143 c      148 g      166 t      3 others
ORIGIN

```

```

Query Match      2.2%; Score 49.8; DB 9; Length 618;
Best Local Similarity 60.9%; Pred. No. 0.0072;
Matches 81; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 722 TCATTGGCGCTGCGATTCGCGAGCTGCTGGTGGCAAGCAAGCTTCATGCTGGGCTAG 781
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 82 TCATTGGCGCTGGAATCTCGGAATATCGGCGGCGAAGATTCTCGTGAAGCGAGTAG 141

Qy 782 ACATGTACAAATATATGAAGCAAGTATGCTTTGAGGCAAGCTTTGTCACAGCTT 841
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 142 AATATGATTGATATCTGGAAGCAAGATCGATCGAGAGAAATACATTAACAGAAATT 201

Qy 842 TCAGGACGCTCC 854
    |||:|||||:|||||:
Db 202 TCGGTACGTGCC 214

```

```

RESULT 3
BU123611      564 bp      mRNA      linear      EST 23-JAN-2002
LOCUS
DEFINITION
BU123611 unpublished oligo-capped cDNA library, C. elegans L1 stage
Caenorhabditis elegans cDNA clone yk1302a06 5', mRNA sequence.
ACCESSION
BU123611
VERSION
BU123611.1 GI:18283750
KEYWORDS
EST.
SOURCE
Caenorhabditis elegans

```

ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Pelodierinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 564)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
source
1..564
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1302a06"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_1ib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"

BASE COUNT 170 a 123 c 115 g 151 t 5 others
ORIGIN

Query Match 2.0%; Score 44.8; DB 12; Length 564;
Best Local Similarity 56.6%; Pred. No. 0.2;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAGTGGCGGTGATGGCGCTGCGACTTCGTGTGCGAAACGACTG 765
DB 244 AAACCGCTATAGCAATCGTCGGCGCGGTATTCGCACTATCGACCGCGCTGACTG 303
QY 766 CTTCATGCTGGGGTACGATGTTACATATATGAAGCAAGTATCTGTGGAGCGCAAG 825
DB 304 ATTGAACCTTGGAATGACGATTTGATATCTACGAAGTCTCGACCGATCGAGCGACANA 363
QY 826 CTTTGGTCACATGCTTTCAGGACG 850
DB 364 ATTCATGCCATACCGTACAGGATG 388

RESULT 4
BJ101485 481 bp mRNA linear EST 18-JAN-2002
LOCUS BJ101485 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1027a04 5', mRNA sequence.
ACCESSION BJ101485
VERSION BJ101485.1 GI:18244155
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Pelodierinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 481)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
source
1..481

/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1027a04"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_1ib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"

BASE COUNT 135 a 107 c 104 g 135 t
ORIGIN

Query Match 1.9%; Score 44.2; DB 12; Length 481;
Best Local Similarity 56.6%; Pred. No. 0.29;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAGTGGCGGTGATGGCGCTGCGACTTCGTGTGCGAAACGACTG 765
DB 264 AAACCGCTATAGCAATCGTCGGCGCGGTATTCGCACTATCGACCGCGGTGACTG 323
QY 766 CTTCATGCTGGGGTACGATGTTACATATATGAAGCAAGTATCTGTGGAGCGCAAG 825
DB 324 ATTGAACCTTGGAATGACGATTTGATATCTACGAAGTCTCGACCGATCGAGCGACAGA 383
QY 826 CTTTGGTCACATGCTTTCAGGACG 850
DB 384 ATTCATGCCATACCGTACAGGATG 408

RESULT 5
BJ103567 486 bp mRNA linear EST 18-JAN-2002
LOCUS BJ103567 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1051f01 5', mRNA sequence.
ACCESSION BJ103567
VERSION BJ103567.1 GI:18246237
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Pelodierinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 486)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
source
1..486
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1051f01"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_1ib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"

BASE COUNT 137 a 104 c 114 g 130 t 1 others
ORIGIN

Query Match 1.9%; Score 44.2; DB 12; Length 486;
Best Local Similarity 56.6%; Pred. No. 0.29;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

0y	706	AAGCGGAAGGCGGCGTATTGGGCGTGGCAATTTCCGAGCTCGTGGTGGCAAGCAACTG	765
Db	193	AAACCGTCTATAGCAATGCTCGGCGCGCGTATTTCCGACTATGACCGCCGCTGACTG	252
0y	766	CTTCATGCTGGGGGTAGACGATGTTACATATATAGAACGACAGTGAATCGTGTGAGGCAAG	825
Db	253	ATTGAACCTGGAATTGACGATTTTGATATCTTACAGAGGCTTCGACCGGATCGAGGCGAGA	312
0y	826	CTTTGGTCAATGCTTTTCAGGGAGC	850
Db	313	ATTGATGCCATACCGTACAGAGATG	337
RESULT 6			
LOCUS	AUI99996		
DEFINITION	AUI99996 unpublished oligo-capped cDNA library, stage L2	mRNA	linear EST 17-JUL-2001
ACCESSION	AUI99996		
VERSION	AUI99996		
KEYWORDS	AUI99996.1 GI:14827576		
SOURCE	EST.		
ORGANISM	Caenorhabditis elegans		
	Caenorhabditis elegans		
	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae		
REFERENCE	1 (bases 1 to 508)		
AUTHORS	Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.		
TITLE	A complementary view of the C.elegans genome		
JOURNAL	Unpublished		
COMMENT	Contact: Yuji Kohara Genome Biology Lab. National Institute of Genetics Yata 111, Mishima, Shizuoka 411, Japan Tel: 81-559-81-6854 Fax: 81-559-81-6855 Email: ykohara@lab.nig.ac.jp.		
FEATURES			
source	1..508	Location/Qualifiers	
	/organism="Caenorhabditis elegans"		
	/mol_type="mRNA"		
	/strain="N2"		
	/db_xref="taxon:6239"		
	/clone="YK774807"		
	/sex="Hermaphrodite"		
	/tissue.type="whole animal"		
	/dev_stage="L2"		
	/clone_lib="unpublished oligo-capped cDNA library, stage L2"		
BASE COUNT			
ORIGIN	148 a	109 c	116 g 135 t
Query Match			
Best Local Similarity	56.6%	Pred. No. 0.3;	Length 508;
Matches	82;	Conservative 0;	Mismatches 63;
		Indels 0;	Gaps 0;
0y	706	AAGCGGAAGTGGCGGCTATTGGCGTGGCAATTTCCGAGCTCGTGGTGGCAAGCAACTG	765
Db	205	AAACCGTCTATAGCAATGCTCGGCGCGCGTATTTCCGACTATGACCGCGGCTGACTG	264
0y	766	CTTCATGCTGGGGGTAGACGATGTTACATATATAGAACGACAGTGAATCGTGTGAGGCAAG	825
Db	265	ATTGAACCTGGAATTGACGATTTTGATATCTTACAGAGGCTTCGACCGGATCGAGGCGAGA	324
0y	826	CTTTGGTCAATGCTTTTCAGGGAGC	850
Db	325	ATTGATGCCATACCGTACAGAGATG	349
RESULT 7			
LOCUS	BUI08542		
DEFINITION	BUI08542 unpublished oligo-capped cDNA library, C. elegans L1 stage	mRNA	linear EST 30-MAY-2003

ACCESSION	Caenorhabditis elegans cDNA clone Yk1109b04 5', mRNA sequence.			
VERSION	B0108542			
KEYWORDS	B0108542.2 GI:31243973			
SOURCE	EST.			
ORGANISM	Caenorhabditis elegans			
	Caenorhabditis elegans			
	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae			
REFERENCE	1 Rhabditidae; Peloderinae; Caenorhabditis.			
AUTHORS	1 (bases 1 to 511) Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.			
TITLE	A complementary view of the C.elegans genome			
JOURNAL	Unpublished			
COMMENT	On Jan 22, 2002 this sequence version replaced gi:18266570. Contact: Tadao Shin-i National Institute of Genetics 111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@gene.nig.ac.jp. Location/Qualifiers			
FEATURES	1. 511			
SOURCE	/organism="Caenorhabditis elegans"			
	/mol_type="mRNA"			
	/strain="N2"			
	/db_xref="taxon:6239"			
	/clone="Yk1109b04"			
	/sex="hermaphrodite"			
	/rilsue_type="whole animal"			
	/dev_stage="L1"			
	/clone_id="unpublished oligo-capped cDNA library, C.			
	elegans L1 stage"			
BASE COUNT	150 a 108 c 117 g 135 t 1 others			
ORIGIN				
Query Match	1.9%; Score 44.2; DB 12; Length 511;			
Best Local Similarity	56.6%; Pred. No. 0.3; 63; Indels 0; Gaps 0;			
Matches	82; Conservative 0; Mismatches			
Oy	706 AAGCGAAATGGCGGCTGATGGCGCTGGGCAATTCGCGATGCGCAAGCAACTG 765			
Db	209 AAACCGCTATAGCAATCGTCGGCGCGATTTCCGCACTATCGACCGGCGCTGACTG 268			
Oy	766 CTTGATCTGGGGTAGACGATGTTACATATATGAGCAAGTAGTCGTGTGGAGGCAAG 825			
Db	269 ATTGAACCTTGGAATTGACGATTTGATATCTACGAAGGTCTCGACCGATCGAGGCGAGA 328			
Oy	826 CTTTGTCACATGCTTCAGGAGC 850			
Db	329 ATTGATGCCATPACCGTACAGGAGT 353			
RESULT 8				
AU203802	513 bp mRNA linear EST 30-MAY-2003			
LOCUS	AU203802 unpublished oligo-capped cDNA library, stage L1			
DEFINITION	Caenorhabditis elegans cDNA clone Yk825e11 5', mRNA sequence.			
ACCESSION	AU203802			
VERSION	AU203802.2 GI:31231885			
KEYWORDS	EST.			
SOURCE	Caenorhabditis elegans			
ORGANISM	Caenorhabditis elegans			
	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae			
REFERENCE	1 Rhabditidae; Peloderinae; Caenorhabditis.			
AUTHORS	1 (bases 1 to 513) Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.			
TITLE	A complementary view of the C.elegans genome			
JOURNAL	Unpublished			
COMMENT	On Jul 17, 2001 this sequence version replaced gi:14834726. Contact: Yuji Kohara Genome Biology Lab.			

National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers

FEATURES

source

```

1..513
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk825e11"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_1lb="unpublished oligo-capped cDNA library, stage
L1"

```

BASE COUNT 141 a 114 c 114 g 143 t 1 others
ORIGIN

Query Match 1.9%; Score 44.2; DB 9; Length 513;
Best Local Similarity 56.6%; Pred. No. 0.3;

Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```

QY 706 AACCCGAAGTGGCGGTCAATTCGCGCTGCGACTCGTGTCGCAACGAACCTG 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 230 AACCGCTATAGCAATCGTCGCGCGCGGTATTTCCGACTATCGACCGCGCTCAGCTG 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 766 CTTTCATCTGGGGTAGACGATGTTACATATATGAAGCAAGTATGTTGGAGCGAAG 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 290 ATTGAACCTTGGAATGACGATTTTGATATCTACGAAGTCTCGACCGAGTGGAGCAGA 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 826 CTTTGTCAATGCTTTCAGGAGC 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 350 ATTCAATCCATACCGTACAGGATG 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 9

BUI20847

LOCUS BUI20847 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1261a07 5', mRNA sequence.

ACCESSION BUI20847
VERSION BUI20847.1 GI:18280980

KEYWORDS EST.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans
Bukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
; Rhabditidae; Pelodierinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 529)
Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
A complementary view of the C.elegans genome

AUTHORS Unpublished
TITLE Unpublished
JOURNAL Contact: Tadasu Shin-i
COMMENT Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES

source

```

1..529
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1261a07"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_1lb="unpublished oligo-capped cDNA library, C.
elegans L1 stage"

```

BASE COUNT 154 a 112 c 121 g 142 t
ORIGIN

Query Match 1.9%; Score 44.2; DB 12; Length 529;
Best Local Similarity 56.6%; Pred. No. 0.3;

Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```

QY 706 AACCCGAAGTGGCGGTCAATTCGCGCTGCGACTCGTGTCGCAACGAACCTG 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 193 AACCGCTATAGCAATCGTCGCGCGCGGTATTTCCGACTATCGACCGCGCTCAGCTG 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 766 CTTTCATCTGGGGTAGACGATGTTACATATATGAAGCAAGTATGTTGGAGCGAAG 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 253 ATTGAACCTTGGAATGACGATTTTGATATCTACGAAGTCTCGACCGAGTGGAGCAGA 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 826 CTTTGTCAATGCTTTCAGGAGC 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 313 ATTCAATCCATACCGTACAGGATG 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 10

AU201678

LOCUS AU201678 540 bp mRNA linear EST 17-JUL-2001
DEFINITION AU201678 unpublished oligo-capped cDNA library, stage L1
Caenorhabditis elegans cDNA clone yk797d07 5', mRNA sequence.

ACCESSION AU201678
VERSION AU201678.1 GI:14830776

KEYWORDS EST.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans
Bukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
; Rhabditidae; Pelodierinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 540)
Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
A complementary view of the C.elegans genome

AUTHORS Unpublished
TITLE Unpublished
JOURNAL Contact: Yuji Kohara
COMMENT Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers

FEATURES
source
1..540
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk797d07"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_1lb="unpublished oligo-capped cDNA library, stage
L1"

BASE COUNT 157 a 117 c 119 g 147 t
ORIGIN

Query Match 1.9%; Score 44.2; DB 9; Length 540;
Best Local Similarity 56.6%; Pred. No. 0.3;

Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```

QY 706 AACCCGAAGTGGCGGTCAATTCGCGCTGCGACTCGTGTCGCAACGAACCTG 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 229 AACCGCTATAGCAATCGTCGCGCGCGGTATTTCCGACTATCGACCGCGCTCAGCTG 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 766 CTTTCATCTGGGGTAGACGATGTTACATATATGAAGCAAGTATGTTGGAGCGAAG 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 289 ATTGAACCTTGGAATGACGATTTTGATATCTACGAAGTCTCGACCGAGTGGAGCAGA 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 826 CTTTGTCAATGCTTTCAGGAGC 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 349 ATTGATGCATACCGTACAGATG 373

RESULT 11

LOCUS CB401025 560 bp mRNA linear EST 15-MAY-2003

DEFINITION OSTF186F11_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.

ACCESSION CB401025

VERSION CB401025.1 GI:30742752

KEYWORDS EST.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

AUTHORS 1. Rhabditidae; Peloderinae; Caenorhabditis.

2. (bases 1 to 560)

Reboul, J., Vaglio, P., Rual, J.F., Lamesch, P., Martinez, M., Armstrong, C.M., Li, S., Jacotot, L., Bertin, N., Janky, R., Moore, T., Hudson, J.R., Hartley, J.L., Brasch, M.A., Vandenhaute, J., Boulton, S., Endress, G.A., Jena, S., Chevet, E., Papasotiropoulos, V., Tollas, P.P., Placsek, J., Snyder, M., Huang, R., Chance, M.R., Lee, H., Doucette-Stamm, L., Hill, D.E. and Vidal, M.

C. elegans ORFome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression

Nat. Genet., (2003) In press

CONTACT: Vidal, M

Journal Comment: Marc Vidal Laboratory Dana Farber Cancer Institute 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA Tel: 617 632 5180 Fax: 617 632 5739 Email: Marc.Vidal@dfci.harvard.edu

Sequence tag of Gateway entry clones. The primers used were designed on the predicted protein encoding ORF. C. elegans ORFome cloning project : Contact david_hill@dfci.harvard.edu or marc_vidal@dfci.harvard.edu

POLYA=No.

FEATURES

source Location/Qualifiers

1..560

/organism="Caenorhabditis elegans"

/mol_type="mRNA"

/strain="N2"

/db_xref="taxon:6239"

/sex="Hermaphrodite and male"

/tissue_type="whole animal"

/dev_stage="mixed stage"

/clone_1lb="AD-wrmcDNA"

/note="The AD-wrmcDNA library was generated with poly(A)+ RNA isolated from both hermaphrodite and male N2 worms of all larval stages, embryos, adults and dauers and the subsequent generation of cDNAs by poly(A) priming. The cDNAs were cloned into pCR6"

BASE COUNT 175 a 116 c 122 g 147 t

ORIGIN

Query Match 1.9%; Score 44.2; DB 14; Length 560;

Best Local Similarity 56.6%; Pred. No. 0.3;

Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAAGTGGCGGTGCTGCGGCTGCGGACTCGTGTGCGAAAGCACTG 765

DB 69 AAACCGCTATAGCAATCGTCGGCGCGGTATTTCCGAGCTATGACCGCGCTGCACTG 128

QY 766 CTCATGCTGGGGTAGACGATGTACATATATGAAACAGATGATCTGTGTGAGGCGAAG 825

DB 129 ATTGAACCTTGGAATGACGATTTGTATATCTACGAAGGCTCTGACCGGATCGAGGCGA 168

QY 826 CTTGTGCACATGCTTCAGGACG 850

DB 189 ATTATGCCATACCGTACAGATG 213

RESULT 12

Bu121613

LOCUS Bu121613 584 bp mRNA linear EST 23-JAN-2002

DEFINITION Bu121613 unpublished oligo-capped cDNA library, C. elegans L1 stage

ACCESSION Caenorhabditis elegans cDNA clone yk1278f01 5', mRNA sequence.

VERSION Bu121613

KEYWORDS EST.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

AUTHORS 1. Rhabditidae; Peloderinae; Caenorhabditis.

2. (bases 1 to 584)

Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.

A complementary view of the C. elegans genome

Unpublished

Journal Comment: Contact: Tadao Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshin@genes.nig.ac.jp.

FEATURES

source Location/Qualifiers

1..584

/organism="Caenorhabditis elegans"

/mol_type="mRNA"

/strain="N2"

/db_xref="taxon:6239"

/clone="yk1278f01"

/sex="hermaphrodite"

/tissue_type="whole animal"

/dev_stage="L1"

/clone_1lb="unpublished oligo-capped cDNA library, C. elegans L1 stage"

BASE COUNT 178 a 122 c 130 g 154 t

ORIGIN

Query Match 1.9%; Score 44.2; DB 12; Length 584;

Best Local Similarity 56.6%; Pred. No. 0.31;

Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAAGTGGCGGTGCTGCGGCTGCGGACTCGTGTGCGAAAGCACTG 765

DB 210 AAACCGCTATAGCAATCGTCGGCGCGGTATTTCCGAGCTATGACCGCGCTGCACTG 269

QY 766 CTCATGCTGGGGTAGACGATGTACATATATGAAACAGATGATCTGTGTGAGGCGAAG 825

DB 270 ATTGAACCTTGGAATGACGATTTGTATATCTACGAAGGCTCTGACCGGATCGAGGCGA 329

QY 826 CTTGTGCACATGCTTCAGGACG 850

DB 330 ATTATGCCATACCGTACAGATG 354

RESULT 13

LOCUS Bu121613 586 bp mRNA linear EST 23-JAN-2002

DEFINITION Bu121613 unpublished oligo-capped cDNA library, C. elegans L1 stage

ACCESSION Caenorhabditis elegans cDNA clone yk1278h02 5', mRNA sequence.

VERSION Bu121613

KEYWORDS EST.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

AUTHORS 1. Rhabditidae; Peloderinae; Caenorhabditis.

2. (bases 1 to 586)

Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.

A complementary view of the C. elegans genome

Unpublished

Journal Comment: Contact: Tadao Shin-i

Mon Nov 24 16:22:33 2003

us-09-434-837-10.rst

Page 8

Db 349 ATTGATGCCATACCGTACAGAGATG 373

Search completed: November 24, 2003, 04:01:39
Job time : 4835.41 secs